From Model, Signal to Knowledge: A Data-Driven Perspective of Fault Detection and Diagnosis

Xuewu Dai, *Member, IEEE*, and Zhiwei Gao, *Senior Member, IEEE*

***Abstract*—This review paper is to give a full picture of fault detection and diagnosis (FDD) in complex systems from the perspective of data processing. As a matter of fact, an FDD system is a data-processing system on the basis of information redundancy, in which the data and human's understanding of the data are two fundamental elements. Human's understanding may be an explicit input–output model representing the relationship among the system's variables. It may also be represented as knowledge implicitly (e.g., the connection weights of a neural network). Therefore, FDD is done through some kind of modeling, signal processing, and intelligence computation. In this paper, a variety of FDD techniques are reviewed within the unified data-processing framework to give a full picture of FDD and achieve a new level of understanding. According to the types of data and how the data are processed, the FDD methods are classified into three categories: model-based online data-driven methods, signal-based methods, and knowledge-based history data-driven methods. An outlook to the possible evolution of FDD in industrial automation, including the hybrid FDD and the emerging networked FDD, are also presented to reveal the future development direction in this field.**

***Index Terms*—Complex systems, data-driven, fault detection and diagnosis (FDD), knowledge-based, model-based, signal-based.**

I. Introduction

INDUSTRIAL systems have been becoming more complex and expensive with less tolerance for performance degradation, productivity decrease, and safety hazards, such as wind farms [28], [98], aircraft engines [20], [53], petrochemical production [89], and metallurgical production [91]. This leads to an ever increasing requirement on reliability and safety of control systems subjected to faults and failures. With the advent of computerized control, communication networks, and information techniques, a huge volume of operation data relating to the process's conditions and status have been collected, which not only makes new fault detection and diagnosis (FDD) methods possible, but also brings challenges.

Manuscript received December 02, 2012; accepted December 21, 2012. Date of publication January 30, 2013; date of current version October 14, 2013. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 60574026 and Grant 61101135. Paper no. TII-12-0815.

X. Dai is with the School of Electronic and Information Engineering, Southwest University, Chongqing 400715, China, (e-mail: daixuewu@swu.edu.cn).

Z. Gao is with the School of Computing, Engineering & Information Sciences, Northumbria University, Newcastle, NE2 1X, U.K. (e-mail: zhiwei.gao@northumbria.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2013.2243743

As an effective means to ensure the reliability and safety of industrial systems and reduce the risk of unplanned breakdowns, FDD has been the subject of interest in control and automation community [13], [42], [51], [52] and finds its success in many engineering areas. FDD studies how to detect the occurrence of a failure as early as possible and how to identify the location and type of the fault as accurately as possible. In the early years, a primitive FDD was simply a limit checker of measurements. Unfortunately, the simple over-threshold checking method becomes invalid as the system complexity increases. Analytical model-based fault-detection methods were proposed to overcome difficulties raised with limit checking. With the mature of state-space modeling and system identification techniques in the 1970s, model-based FDD has become the main stream of research since the 1980s. The model-based method involves rigorous development of process models either derived from first principles or identified from measured data. The representative work of model-based FDD includes parameter identification method, observer-based method, and parity space method. At nearly the same period, the signal-based FDD method was developed due to the significant improvement of digital signal-processing techniques. One of the most successful applications of signal-based FDD is the motor current signature analysis (MCSA) for electric motors and generators.

Recently, with the rapid development of smart instruments, digital communication networks and computer techniques, distributed control systems (DCSs) have been widely deployed in advanced industrial systems and provided the ability to collect and store a huge amount of process data. The emerging DCSs and networked control systems (NCSs) make the data acquisition much easier. The amount of the collected data, however, is too much to be fully and effectively utilized by most existing FDD methods. As a result, "large volumes of data with very little information" is a quite common phenomenon in today's industrial automation. For instance, in the condition monitoring of wind farms, there are a number of various databases with data and statistics, but it is difficult to get an overall picture of the relationship between failures and data [81]. Enabled by the ever increasing computational power governed by Moore's law, many artificial intelligence (AI) techniques in computer science have been introduced to FDD to deal with the huge amount of data and extract *useful information* (or termed by knowledge) from data [8], [63]. Particularly, in the 1990s, *machine learning* (sometimes referred to as *soft computing* or *computing intelligence* [55], [97]) was developed, which mimics human's abilities of logic reasoning by numeric computing and connections, rather than by the traditional logic algebra developed in the 1950s. Typical examples of soft computing are neural networks and fuzzy logics [55]. The introduction of computing intelligence develops a new trend of knowledge-based FDD methods [86].

The new trend in FDD is to integrate various strategies to form a hierarchical structure with a mixture of various homogeneous and/or heterogeneous FDD methods. Consequently, the study of FDD has been a multidisciplinary field involving control engineering, signal processing, and artificial intelligence (AI). The diversity of the FDD methods makes it difficult for an engineer to master all of the techniques and trends in different fields. In particular, it seems that results from AI play, and will continue to play, an important role in FDD. It is necessary to find their common features and difference and build a systematic view to represent the new trends in FDD under a unified framework.

Nevertheless, the fact is that any FDD system is a system of data processing on the basis of information redundancy [14], [35], in which the data itself and the understanding of the data are two fundamental elements. Different FDD strategies vary at how the data are understood and how the information behind the data is exploited. In this survey, with the purpose of providing a full picture of FDD including these signal-based, model-based, and knowledge-based approaches, we study these strategies from the viewpoint of how the data are processed for FDD. This is a systematic and comparative study of various FDD strategies by examining the relationship among information, data, model, signal, and knowledge under the data-driven framework. We attempt to present a data-driven perspective showing how these different methods relate to and differ from each other.

The remainder of this paper is organized as follows. As a preparation, Section II examines the relationship among data, models, signals, and knowledge in FDD. Section III reviews the model-based online-data-driven FDD followed by signal-based FDD in Section IV. In Section V, knowledge-based history-data-driven FDD is investigated. Section VI presents an outlook of the possible evolution of FDD in advanced industrial automation. We end the paper with a conclusion in Section VII.

II. Categories of FDD

Here, we start from the viewpoint of information redundancy and data-driven FDD, where an FDD always makes use of data and models either explicitly or implicitly. We then classify FDD into three categories, investigate the core concepts in these categories, and study their relationship.

In industrial automation, FDD is used to monitor the behavior of a process that is usually described as a dynamic system. Here, a dynamic system is a process producing outputs from inputs, in which variables of different kinds interact, and the output variables depend on the present and past values of the input variables. From the viewpoint of information theory, the correlation and dependences among these variables are information redundancy, which is the basis of all FDD. A traditional approach to have information redundancy is physical redundancy that is the duplication of hardware components (e.g., controllers and sensors). Another form of redundancy is analytical redundancy, in which the correlation among the related variables are represented either explicitly by a mathematic model or hidden behind the huge amount of data in an implicit form.

Since most FDD algorithms nowadays are carried out by digital processors in the discrete-time domain on the basis of sampled data, only discrete systems are included herein. Consider a system with m inputs (denoted by $\mathbf{u}[k] = [u_1[k] \cdots u_m[k]]^T$) and p outputs (denoted by $\mathbf{y}[k] = [y_1[k] \cdots y_p[k]]^T$), where k is the discrete time, the relationship between $\mathbf{y}[k]$ and $\mathbf{u}[k]$ is written as a function

$$\mathbf{y}[k] = F(A(z)\mathbf{y}(k), B(z)\mathbf{u}[k], \theta) \quad (1)$$

where $A(z)$ and $B(z)$ are polynomial with respect to the backward shift operator z^{-1}, θ is the systems parameters.

In (1), the known function F represents the analytical redundancy explicitly. When the dynamic system gets more complex, it becomes impossible to have such an explicit function. Defining the measurements of variables as signal or data and referring the implicit dependency behind data as knowledge, we can tell if the dynamic system has faults by checking consistence between the data and knowledge. The data should match with the expected knowledge if the system works in good condition as expected. In this sense, knowledge and data are redundant to some extent.

In the context of information redundancy, an analytical FDD is a data (signal) processing with one search engine to check information redundancy between the data and explicit model or implicit knowledge. Here, redundancy checking means to check the consistency of the data against a model or knowledge or to directly check the consistence among the data themselves. In this sense, FDD methods are always data-driven on the base of model or knowledge.

In this paper, we investigate the analytic FDD methods from the viewpoint of how the data are processed for FDD. Depending on how the data and the dependency are deployed, FDD methods can be classified into three categories, namely model-based (online-data-driven) FDD, signal-based (data-driven) FDD, and knowledge-based (history-data-driven) FDD. This concept is illustrated in Fig. 1 schematically.

The bottom of Fig. 1 depicts the model-based FDD, in which only a small amount of online data is used to detect and diagnose faults. A mathematic model M with parameter θ has been available from first principles or identified through system identification techniques. The data of system input and output are then fed into the data-processing engine that generates residuals by comparing the measured data and model's predictions. A residual classifier or θ classifier is next employed to check if there is a fault and decide what fault it could be. A good model-based FDD ideally has residuals sensitive only to system faults but not to disturbances or deviations in system inputs (such as motor power supply imbalance or motor load variations).

The block diagram of the signal-based FDD is shown in the middle of Fig. 1. The information redundancy in signal-based FDD methods is the relationship between faults and the signal patterns. Since the faults within the system usually have direct influences on output variable y_k, it is straightforward that the signal used in most signal-based FDD methods is the sampled

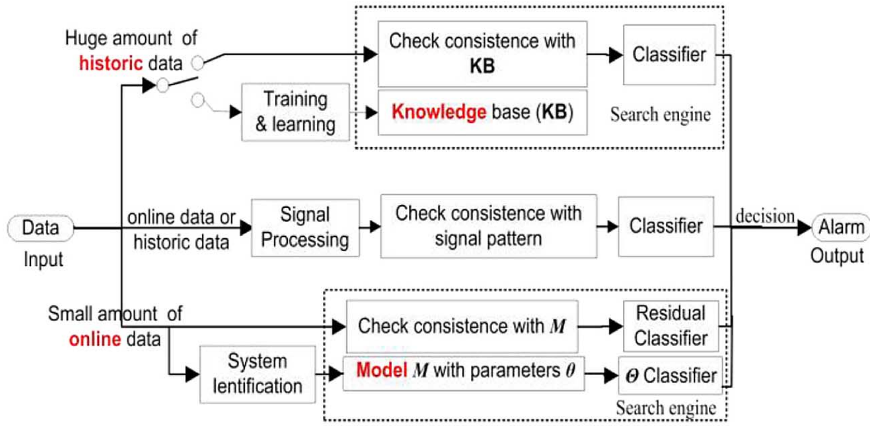

Fig. 1. Data flow in FDD.

output variable y_k and there is no need for an input-output model of the dynamic system. This is beneficial for complex industrial process or machine systems where accurate input-output models are usually unavailable and/or their parameters are hard to estimate.

When a process is too complex to be modeled analytically and the signal analysis does not yield an unambiguous diagnosis, a sophistic FDD approach aided by expert systems or AI has to be used, which usually involves a huge amount of history data. This leads to the knowledge-based (historic) data-driven FDD, of which the block diagram is shown at the top of Fig. 1. In a narrow sense, the knowledge-based FDD is often referred to as "data-driven" FDD, since it is very common in a complex industrial process that only a huge amount of data is provided and the explicit models or signal patterns of a system are not available straightforward. Such a data-driven FDD is based on the implicit knowledge mined from the huge amount of history data through some intelligent training or machine learning methods. Once the knowledge is developed from the history data to form a knowledge-base (KB) implicitly representing the dependency of the system's variables, the consistency between the recent data and the KB is checked and a classifier follows to make final decision.

In signal-based FDD, the relationship between output signal y_k and faults are built up from human's *a priori* understanding of the system. On the other hand, knowledge-based FDD discovers the dependency from a huge amount of data autonomously. This is a distinguishing character of knowledge-based FDD different from signal-based FDD. The measured signal y_k possess some features in the time domain and/or frequency domain, which can, in many cases, be mean, variance, frequency, magnitude, and phase. Most importantly, these features are linked to the faults. Different faults result in different combinations of these features, and the combination of features is referred to as signal pattern or signal signature. Obviously, different faults have related signals to show different patterns. As a result, the basic data processing in signal-based FDD is to extract the features from the signal to get their patterns and compare the signal pattern with known pattern to detect and diagnose faults. Depending on the signal-processing techniques (statistical or nonstatistical) and the patterns used in FDD, the data required for signal-based FDD can be online data or history data. In most cases, the data size in signal-based FDD is larger than the model-based method, but much smaller than in the knowledge-based method.

Fig. 1 also shows that a FDD has three elements: 1) a representation of information redundancy, which can be explicit mathematic models, known signal patterns, or implicit KB derived from data; 2) data collected during operation which will be checked against the information redundancy; and 3) a consistence-check engine with classifiers. The consistence-check engine in turn depends on the type of data available and the form of information redundancy. From this point of view, we classify the data-driven FDDs into these three categories according to the type of data and form of information redundancy.

In Sections III–V, these three categories of data-driven FDD techniques will be reviewed, respectively.

III. Model-Based Online Data Driven FDD

The model-based FDD methods have been fruitful and, for the sake of analysis, the input–output model $F(\cdot)$ of system (1) is transformed into a general state-space model

$$\mathbf{y}_k = F(\mathbf{u}_k, \mathbf{x}_k, \mathbf{d}_k, \mathbf{f}_k, \theta) \tag{2}$$

where subscript $_k$ denotes time index, $\mathbf{x}_k \in \mathbb{R}^n$ is an n-dimensional state vector, $\mathbf{d}_k$ is the unknown input denoting modelling errors, measurement noises, and external disturbances, and $\mathbf{f}_k$ represents possible faults to be detected.

Since faults usually cause changes in state variable $\mathbf{x}$, in model parameters θ, and/or have output y derivate from expected values, one can check these changes/derivations to tell if the system has a fault. Based on the explicit model (2), the model-based FDD methods generate output estimates $\hat{y}$, parameter estimates $\hat{\theta}$, and/or state estimates $\hat{x}$ from the data pair $\{\mathbf{u}_k, \mathbf{y}_k\}$. Checking these estimates with respect to their expected nominal values, a residual $\mathbf{r}$ is generated and classified. Accordingly, model-based FDD consists of three main branches: 1) parameter estimation method resting from system identification [51]; 2) parity relation approach [14], [42]; and 3) observer/filter-based approach [34].

A. Parameter Estimation for FDD

In most applications, the parameters are unmeasurable, but they can be determined with parameter estimation methods from measured input/output data $\{\mathbf{u}_k, \mathbf{y}_k\}$. The parameter estimation methods have been extensively studied in system identification [64], and its application to FDD was first described by [51] as follows. The model's parameters $\theta = f(\varphi)$ are related to physically defined process coefficients φ (like resistance, stiffness, and loads). Faults within the system will have a change $\Delta\varphi$ in φ. When θ is estimated, and, in turn, $\Delta\varphi$ is computed by solving $\theta = f(\varphi)$ and fault can be detected and diagnosed. Hence, the FDD problem turns into parameter estimation, which can be solved by least-square error (LSE) and its derived methods [64], such as instrumental variables and (recursive) subspace methods [18]. Various parameter-estimation methods for FDD are reviewed in [51] and [52]. Regardless of the parameter-estimation methods employed, the logics of FDD are the same as those suggested in [51].

A high-gain observer-based online parameter-estimation method was recently proposed in [40] for a system subject to bounded process and measurement noises. In this approach, the parameter changes are modeled as an unknown disturbance $d(t)$. A high gain observer is then applied to estimate $d(t)$, and a linear square estimation method is applied to estimate the parameter changes from $d(t)$.

The main advantages of parameter identification-based method are that the fault diagnosis is very straightforward if the model parameter has a one-to-one mapping with the physical coefficients. For example, function $f(\varphi)$ is an identity matrix or the model is a gray-box model. Detecting sensor/actuator faults by parameter identification may be complicated, as sensor/actuator faults may influence the input/output in the same way as the process (parameter) faults.

B. Observers and Filters for FDD

The Kalman filter and Luenberger observer based methods have been widely accepted for state estimation and residual generation [13], [75]. For illustration purposes, we consider system in (2) as a linear state-state space model

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{B}_d\mathbf{d}_k + \mathbf{B}_f\mathbf{f}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{D}_d\mathbf{d}_k + \mathbf{D}_f\mathbf{f}_k. \end{cases} \tag{3}$$

For nonlinear systems, the basic steps and concepts are similar, but with nonlinear observer or extended Kalman filter rather than linear ones. The observer (filter) for system (3) is

$$\begin{cases} \hat{\mathbf{x}}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k + \mathbf{K}(\mathbf{y}_k - \hat{\mathbf{y}}_k) \\ \hat{\mathbf{y}}_k = \mathbf{C}\hat{\mathbf{x}}_k + \mathbf{D}\mathbf{u}_k \end{cases} \tag{4}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the estimates of the state and output, respectively, and $\mathbf{K}$ is the *observer gain* to be designed. The diagram of the observer is illustrated in Fig. 2(a). Let e_k denote the state estimation error ($e_k = x_k - \hat{x}_k$) and r_k denote the output estimation error ($r_k = y_k - \hat{y}_k$), the dynamics of the observer (4) are governed by

$$\begin{cases} \mathbf{e}_{k+1} = (\mathbf{A} - K\mathbf{C})\mathbf{e}_k + (\mathbf{B}_d - K\mathbf{D}_d)\mathbf{d}_k + (\mathbf{B}_f - K\mathbf{D}_f)\mathbf{f}_k \\ \mathbf{r}_k = \mathbf{C}\mathbf{e}_k + \mathbf{D}_d\mathbf{d}_k + \mathbf{D}_f\mathbf{f}_k \end{cases}. \tag{5}$$

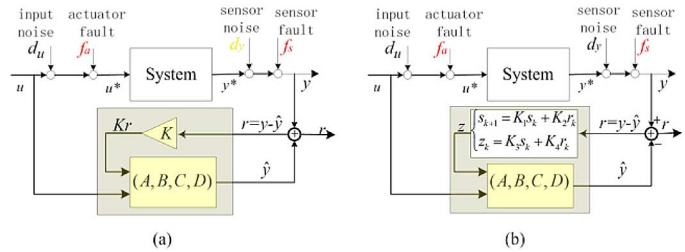

Fig. 2. (a) Static observer and (b) Dynamic observer for FDD.

Applying the Z-transform to (5), the relationship from d_k, f_k to residual r_k in the z-domain is

$$r(z) = \bar{G}_f(z)f(z) + \bar{G}_d(z)d(z) \tag{6}$$

where the transfer function matrices are

$$\begin{cases} \bar{G}_f(z) = C(zI - A + KC)^{-1}(B_f - KD_f) + D_f \\ \bar{G}_d(z) = C(zI - A + KC)^{-1}(B_d - KD_d) + D_d. \end{cases} \tag{7}$$

Equation (6) suggests that the residual r_k is related to both the faults and disturbances. The heart of the observer-based FDD is to make r_k sensitive to faults f_k but insensitive (robust) to disturbance d_k. As one of the dominant FDD approaches, the de-coupling approach has been developed in last two decades [13], in which the disturbances and model uncertainties are treated as unknown inputs and de-coupled from residuals. Unknown input observers (UIOs) were first employed in [95], where the insensitivity to disturbances was achieved indirectly by making the state-estimation error de-coupled from d_k. The direct UIO decoupling r_k from d_k was proposed in [69] and [95] by using eigenstructure assignment. However, the perfect de-coupling may not be possible, when required sufficient condition is not met [13]. An approximate de-coupling should be taken, where the residual is not perfectly de-coupled from disturbances, but has a low sensitivity to disturbances and high sensitivity to faults. It becomes an optimization problem and has been studied both in the time and frequency domains. Some researchers applied multi-objective optimization to solve this problem [12], [33]. To address the nonlinearity of complex systems, sliding-mode observers were developed for fault detection in [29]. H_∞ optimization and linear matrix inequality (LMI) for robust residual generation have received more attention recently [54].

Note that the gain $\mathbf{K}$ of the observer (4) is a numerical matrix which simply amplifies r by K when feeding back the observed information to update the observer. Since the frequency response of the feedback path is a constant value K over all frequencies, the feedback gain does not change the frequency shape of r selectively. This kind of observer is termed as *static observer*, of which zeros are invariant [21]. In order to improve the observer's frequency response, the concept of *dynamic observer* was developed, and a joint pole-zero assignment was proposed in [21], where the numerical gain matrix K is replaced with a dynamic system

$$\begin{cases} s_{k+1} = \mathbf{K}_1 s_k + \mathbf{K}_2 r_k \\ z_k = \mathbf{K}_3 s_k + \mathbf{K}_4 r_k. \end{cases} \tag{8}$$

Fig. 2(b) illustrates the structure of the dynamic observer. By introducing a dynamic system into the observer's feedback path, the observer has some desired ability of frequency shaping to improve the residual's robustness against the disturbances but keep the information of faults.

Another branch of observer-based FDD is fault estimation, including proportional integral observer [38], [85], adaptive observer [92], and sliding-mode observer [88]. They were developed for fault diagnosis and fault-tolerance control and have the ability to estimate the actuator/sensor faults. For a system subjected to input noises and sensor noises, it is more challenging to estimate the fault. Another approach is the descriptor observer [37], [39], where derivative gain is tuned to attenuate sensor noises and high-gain proportional gains to attenuate process noises.

C. Parity Equations for FDD

Another main approach in model-based FDD is parity equations. The data process in parity methods is to check the parity (consistency) of the models with sensor output and known inputs. The idea of parity space approaches can be explained as follows [14], [52]. Consider the state-space model (3), after observing β pairs of input–output data $\{\mathbf{u}_{k+j}, \mathbf{y}_{k+j}\}, 0 \leq j \leq \beta$, the input–output relationship can be rearranged into a compact form

$$\mathbf{Y}_k = \mathbf{T}\mathbf{x}_{k-\beta} + \mathbf{Q}\mathbf{U}_k \tag{9}$$

where $\mathbf{T}$ and $\mathbf{Q}$ are defined as [52]. Left-multiplying (9) with a vector $\mathbf{w}^T$ gives a scalar equation

$$\mathbf{w}^T\mathbf{Y}_k = \mathbf{w}^T\mathbf{T}\mathbf{x}_{k-\beta} + \mathbf{w}^T\mathbf{Q}\mathbf{U}_k. \tag{10}$$

When the state variables $\mathbf{x}_{k-\beta}$ is eliminated, (10) becomes a parity equation, and the residual is generated as

$$\mathbf{r}_k = \mathbf{w}^T(\mathbf{Y}_k - \mathbf{Q}\mathbf{U}_k). \tag{11}$$

Eliminating the state variables $\mathbf{x}_{k-\beta}$ requires $\mathbf{w}^T\mathbf{Q} = 0$, which can be solved if the system is observable. Under healthy conditions, the residual $\mathbf{r}_k$ of the parity equations is zero. Dynamic parity relations were studied in [14] and significantly developed in [42].

There have been many survey papers for model-based FDD [34], [35], [51], [89] by Isermann, Patton, Frank, and Ding, respectively. Recent books like [13] and [52] provide a comprehensive overview of model-based FDD, which are good references for further readings.

IV. Signal-Based Data-Driven FDD

Signal-based FDD is based on analysis of the output signals $y[k]$ and does not involves an explicit input-output model $F(\cdot)$ of the target system. As shown in (2), the system output depends on the system parameters θ. Since a fault within the system usually makes θ deviate from its nominal value, the system's output will change accordingly. More specifically, the pattern and features of the system output signal usually have correlation with faults. Such correlation is the basis of signal-based FDD. Thus, one can monitor and analyze the output signals and find their feature patterns and links to faults, which will provide useful indication of the faults and their types.

Typical signals include vibration, speed, force, current, and magnetic flux density. Even though thermal and other signals have been utilized in FDD, signal-based FDD methods are particularly interesting for motors and rotary machines and mainly focuses on electronic signals and vibrations. The overwhelming majority of motor FDD systems use motor measurements, such as motor currents, negative sequence currents, and/or vibration levels.

Features of the monitored signals are extracted to analyze its patterns, which can be in time and/or frequency domains. Examples of features are signal means, variance, trends, instantaneous power fast Fourier transform (FFT), or the spectra in a frequency band of interest. Typical signal analysis techniques include FFT, spectral estimation, wavelet transform [5], and sequence analysis [71]. Moreover, parametric signal models (e.g., an ARMA model) can be used [52], which allow the main frequencies and their amplitudes to be directly estimated. This approach is especially sensitive to small frequency changes.

Depending on the types of signal patterns and signal analysis techniques, the signal-based FDD methods can be classified into three categories: time-domain, frequency-domain, and joint-time-frequency methods.

A. Time-Domain Signal-Based FDD

It is straightforward to regard a signal as a time-domain waveform and a signal with many characteristics in the time domain, such as period, peak, mean, and standard deviation [13]. Higher order statistics such as root-mean-square (rms), skewness, and kurtosis and crest factor have been used as well [12].

Cross-correlation analysis is a widely accepted technique in the time domain for fault detection and classification. The cross-correlation coefficient $r_{xg} = (\mathrm{cov}_{xy}/\sigma_x\sigma_y)$ provides a dimensionless measurement of linear dependency between two signals x and y. For fault detection and classification, a set of baseline signals in various known conditions are first collected as $\{x\}$, and the correlation analysis between the signal y to be monitored and the baseline signals $\{x\}$ are carried out. The resulting correlation coefficient r_{xy} indicates the possibility of the present condition is. If $|r_{xy}|$ approaches 1, it is highly possible that the system is in the condition corresponding to x. If xy is around zero, the system is not in the condition associated with x.

The negative log-likelihood value is recently proposed for vibration signal-based FDD of mechanic systems [87]. The Weibull negative log-likelihood value (Wnl) and the normal negative log-likelihood value (Nnl) of the time-domain signals are statistical features, which represents the likelihood of the signal's distribution. Combined with neural network classifier, the introduction of Wnl and Nnl benefits fewer input features to neural network and was demonstrated the potential suitability for detecting bearing faults [87].

Most signal-based FDD treat the signal in one-dimensional (1-D) time domain. Recently, an interesting time signal to a 2-D image translation approach is demonstrated in [24]. As illustrated in Fig. 3, the magnitudes of $\boldsymbol{L}$ data samples in a time series are treated as pixel intensity, and the data are rearranged into an $m \times n$ gray image $(m \times n = L)$. The features of the image are extracted through a scale-invariant feature transform

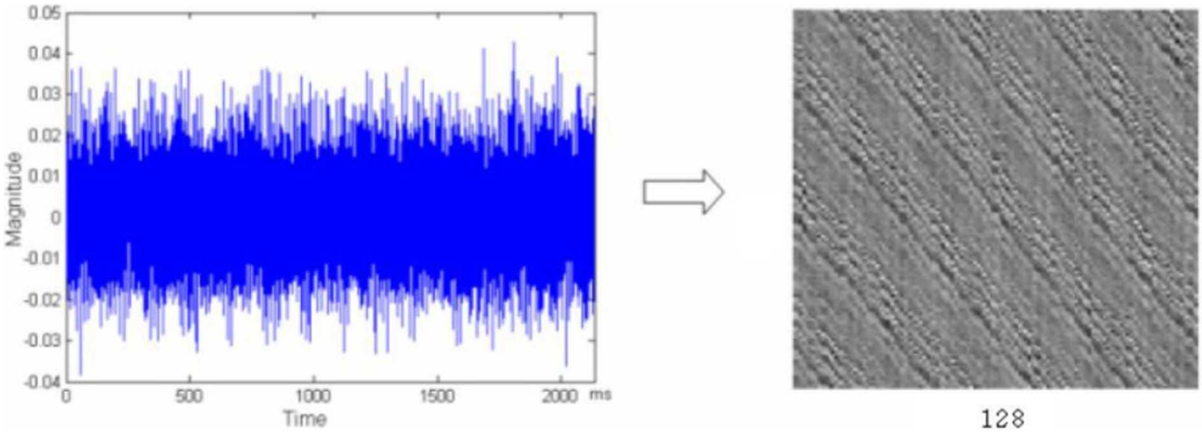

Fig. 3. Example of a 16-k vibration signal translated into a 128×128 gray image [24].

(SIFT) is applied to the image to extract the 2-D local features, which are correlated to faulty symptoms. Therefore, faults can be detected and diagnosed by using advanced image processing and pattern recognition algorithms.

B. Frequency-Domain Signal-Based FDD

Signals related to many mechanical and electrical faults contain feature frequency components and different faults may result in different signal pattern in frequency domain. In most cases, these frequencies can be determined from *a priori* knowledge or known parameters of the target system, e.g., the number of poles of a motor. The use of the frequency analysis of vibration and current signals has been heavily researched to detect bearing, stator, rotor, and eccentricity faults.

Frequency-domain analysis begins by converting a time-domain waveform into its frequency-domain equivalence and the discrete Fourier transformation (DFT) is the most common method used for online condition monitoring. Since the pattern of the dominant frequency components is likely to be the signature of fault, when the frequency spectrum is available, peak detection can be used to identify the dominant frequencies, and envelope analysis [17] can be utilized to describe the patterns including the spacing of sidebands and the presence of harmonics. Silva *et al.* [17] obtained the envelope by using sampled positive peaks of the stator current and extracted fault signature from the envelope using a statistical clustering technique called Gaussian mixture model (GMM). The signature was then fed to a maximum-likelihood Bayesian classifier for diagnosis, which was found to be 99% accurate in detecting a single turn short under 50% rated load.

As one of the most successful signal-based FDDs, the motor current signature analysis (MCSA) has been widely used in modern industrial drive systems [71]. Recent development in MCSA is motor fault detection under unbalanced conditions [4], [16] and condition monitoring of wind generators, such as doubly-fed induction generator (DFIG) [28]. In [2], an adaptive algorithm for fault detection in DFIG was proposed for FDD under dynamic conditions. The work in [4] and [16] studied the multiple reference frames theory that was shown to be immune to voltage unbalances or nonstationary conditions. In [28], the experiment results validated the theory analysis that the current spectrum for a 30-kW DFIG with one broken bar showing the characteristic 1 ± 2 s broken bar sidebands around the 50-Hz peak. This frequency pattern can be employed to detection broken bar faults in wind generators.

C. Joint Time-Frequency-Domain Signal-Based FDD

Individual features in either the time or frequency domains are generally unable to extract all underlying signal information. Time-frequency analysis combines both the time-domain waveform and the corresponding frequency spectrum. This enables the examination of transient features, such as impacts and fault events, as well the ability to monitor frequency content over time [90].

The short-time Fourier transform (STFT) is a common technique, where the signal is divided up into short-time segments, and then a FFT is applied to each window. The Wigner–Ville distribution (WVD) overcomes this resolution limitation in STFT, but it suffers from interference terms forced by the transform itself. Improved transforms, such as Choi–Willams distribution, Zao-Atlas-Marks (ZAM) distribution and cone-shaped distribution, have been developed to further advance time-frequency analysis [79]. In [93], STFT, wavelet transform and the pseudo-Wigner–Ville distribution are investigated for condition diagnosis of rotating machinery. In [3], the stochastic subspace-based identification method was developed.

The trend in signal-based FDD is moving towards application of nontraditional computational techniques in the subject areas such as finite elements and, more recently, wavelet signal processing that has been receiving much attention [5], [15]. For the purpose of analysis, consider 1-D continuous wavelet transform (CWT) of signal $x(t)$ given by

$$X_{WT}(t,s) = \frac{1}{\sqrt{|s|}} \int x(\tau) \cdot \psi^* \left(\frac{\tau - t}{s} \right) d\tau$$

where s is the scaling factor and $\psi^*(\tau - t/s)$ is the basis wavelet function. Different from STFT, the wavelet transform uses scalable basis function $\psi^*()$ and variable size windows, allowing for the acquisition of multiscale resolutions [20]. The discrete wavelet transform (DWT) has also received praise for its computation efficiency and ability to reduce noise in raw signals [12]. DWT has been performed on the vibration and motor current signals and various basis wavelet functions have been proposed for FDD, such as Gaussian-enveloped oscillation wavelet [15], Daubechies family, Symlets family [9], and B-spline (FBS) wavelets that enable an efficient filtering in the region neighboring the main frequency, as well as enable a high level of detail in the time–frequency maps [45]. Discrete wavelet packet transform (DWPT) was proposed to enhance the power and the flexibility of the DWT [9]. Various adaptive methods have been proposed for the selection of optimal basis wavelets [9], [99].

Although it has been demonstrated that these three signal-based approaches are able to work individually to detection and diagnose faults, there have been many reports in literature that combine these methods together. For instance, in [5] and [93], the wavelet analysis and MCSA are integrated. More recent development of the hybrid FDD methods will be discussed in Section VI-B.

V. Knowledge-Based Historic Data-Driven FDD

For those systems which are too complicated to have an explicit system model or signal symptoms, a learn-by-example mechanism is desirable to automate FDD. In contrast to the model/signal-based FDD which requires *a priori* known models or signal patterns, the knowledge-based FDD starts from where only a large amount of historic data is available. Enabled by the advanced artificial intelligence, the knowledge-based FDD learns from empirical data to "*discover*" the underlying knowledge that represents the information redundancy among the system's variables. The intelligent learning from a vast volume of data is the definition feature distinguishing knowledge-based FDD from model-based and signal-based ones, as the latter only require a small amount of data for redundancy checking rather than redundancy learning. Due to this fact, knowledge-based FDD has been commonly referred to as "*data-driven*" FDD, and this name has been widely accepted. However, the term "data-driven" is confusing and less rigorous, as every FDD methods, including model-based and signal-based ones, is a data-processing procedure driven by data. In this paper, it is more scientific to use the full name *knowledge-based historic-data-driven* FDD or shortly *knowledge-based* FDD.

The knowledge-based FDD has become a hot interdisciplinary research topic in the last decade, due to the rapid development of machine learning (ML) in artificial intelligence (AI) since the 1990s. It can be seen that these newly proposed intelligent FDD methods are always lighted by new techniques developed in AI. Because of the close links between knowledge-based FDD and AI, in order to give reader a full picture of the knowledge-based FDD and its trend, it is helpful for such a survey paper to first review the links between AI and FDD briefly followed by detailed discussion on various knowledge-based FDD techniques.

A. AI and Machine Learning in FDD

The knowledge in FDD can be either quantitative or qualitative and is usually organized as a knowledge-base (KB). The KB can be in very different forms, for example, the fault tree is a typical qualitative KB, and a neural network with weighted links forms a quantitative KB. On the other hand, as a knowledge development and management method, AI has adopted two main paradigms: symbolic intelligence and connectionist intelligence. The first is based on symbolic algebra to manipulate symbols. The second is also referred to as computational intelligence, as it is based on computation-intensive machine learning techniques. These two paradigms are associated to qualitative knowledge and quantitative knowledge, respectively.

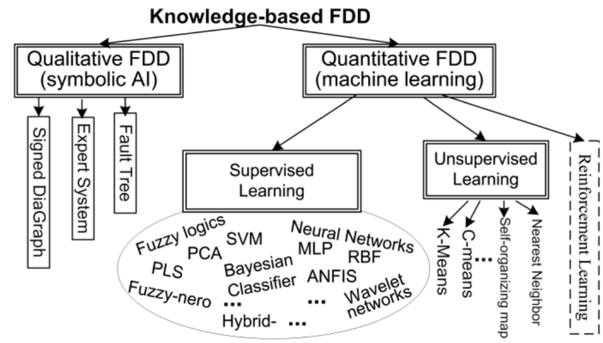

Fig. 4. Knowledge-based history-data driven FDD.

Consequently, it is intuitive to group the knowledge-based FDD into two groups (as shown in Fig. 4): qualitative methods on the basis of symbolic intelligence and quantitative methods on the basis of machine learning intelligence.

The qualitative methods include three subcategories: fault tree (FT), signed diagraph (SDG), and expert system (ES). FT originally developed at Bell Lab in the 1960s is a logic cause–effect tree that propagates primary events (faults) from bottom to the top level events (symptoms). A recent application of FTs in FDD was reported in [57] for reliability analysis and fault diagnosis. SGD is a graph with directed arcs leading from a "cause" node to "effect" nodes, and these arcs are given a positive or negative sign. SDG have been the most widely used form of qualitative knowledge in FDD. ES is generally a tailored system containing deep, but in a narrow-domain expertise of a system. The expert system indeed is a rule-based system presenting human's expertise in a set of rules. Initial attempts at the application of expert systems to fault diagnosis can be found in [73]. In [100], a methodology was presented for formulating diagnostic rules from the knowledge of system structures and component functions. A fuzzy expert system was proposed in [30], and interested readers should refer to [61].

These qualitative FDD are based on the traditional symbolic AI that was first developed in the 1950–1960s and revived in the 1980s due to the success application of experts system in condition monitoring. Nowadays, enabled by the exponentially increasing computation power, computational intelligence (also called machine learning or "soft computing" [55]) has become the most attractive AI techniques. As the ML is an effective way to obtain knowledge from a huge amount of empirical data at the cost of intensive computation, it is straightforward to apply ML for detecting and diagnosing faults from data without the need for explicit model.

Fig. 4 shows a schematic classification of the quantitative knowledge-based FDD from the viewpoint of machine learning.

It is noticed that, in these quantitative knowledge-based FDD, the history data is first transformed by ML into knowledge. This procedure is known as *training* or *learning*. Since the dominant machine learning techniques used in FDD are unsupervised learning and supervised learning, we only discuss these two methods in this paper.

B. Supervised Learning for FDD

In supervised learning FDD, the data is first classified and labeled with tags that indicate the system's conditions and symptoms, such as healthy, faulty and the type of faults. The labels are also known to the machine learner. Here, by 'machine learner' we mean the machine learning algorithms. The machine learner's task is to search for patterns and rules representing the information redundancy and relationship between data patterns and faults. Typical machine learner in knowledge-based FDD are neural nework, fuzzy logics, and PCA, etc.

1) Neural Networks (NNs): NNs are one of most well-established machine learning techniques for monitoring complex nonlinear processes. An NN is a set of nodes linked by connections with weights representing the "strength" of those connections. The nodes are organized into layers and data is propagated through successive layers. The input-output relationship of ith node at the jth layer is a nonlinear function

$$y_i^j = f_i^j(\cdot) = f_i^j\left(\sum_{k=1}^{N} w_{k,j-1}^{i,j} y_k^{j-1}, \theta_i^j\right) \tag{12}$$

where y_i^j is the output of the ith node at the jth layer, w_k^j is the connection weight from the kth node at the $(j-1)$th layer to the ith node at the jth layer, $\boldsymbol{N}$ is the number of inputs (usually equal to the number of preceding nodes), and θ_i^j is the node's parameter. It can be seen that the overall function of NNs is a series of superposition and composite function of $f_i^j(\cdot)$. The most common $f_i^j(\cdot)$ is the sigmoid transfer function $f(x) = 1/(1+e^{-x})$ or a (Gaussian) radial basis function $f(x) = \exp[-\beta\|x - \mathbf{c}\|^2]$.

In FDD, the input to the NN is the history data set and the final output is an indication of the target system's status (healthy or faulty). Given the dimension of the data set is m and the number of possible type of faults is n, the relationship can be expressed by a m-to-$(n+1)$ function $G: m \rightarrow n+1$ mapping from m-dimensional data to $(n+1)$-dimensional health/fault status. Due to the complexity of the target system, function G is usually very complicated and highly nonlinear, and getting an analytic form of G is extremely difficult or impossible. Since NNs have shown its good ability to approximate complex nonlinear functions, it is feasible and straightforward to use an NN to approximate G. The most important stage in NN-based FDD is training, in which the connection weights w_k^i and node's parameters θ_i^j are adjusted by some training algorithm to have the NN approximate G. More specifically, the training is an optimization process to minimize the approximation error between NN and the desired function G. The most popular supervised learning strategy in NNs is back-propagation algorithm [55], [67].

Due to its powerful nonlinear function approximation and adaptive learning capabilities, NNs have drawn great attention in FDD. In chemical engineering, one pilot study of neural networks for FDD was reported in [49]. The NN method was later extended to utilize dynamic process data [86].

Most of the work on improvement of NNs for FDD is based on the selection and modification of function $f(\cdot)$. References [56] and [72] suggested the use of a radial basis function for FDD. In [59], the radial function was extended to Gaussian functions and the hidden node problem was addressed for large-scale fault diagnosis.

Different network architectures have also been proposed for FDD [44]. NNs are also integrated with other machine learning algorithms to improve the fault diagnosis performance. A very common one is the combination of fuzzy logics with neural networks. In [55], a typical fuzzy-neural network was proposed and a number of successful applications can be found in [11].

2) Fuzzy Logic (FZ): FZ is a means of partitioning a feature space into fuzzy classes and using fuzzy rules for reasoning. In contrast to NNs in which the knowledge is implicitly represented by a network of connections implicitly, FZ has the advantages of describing human knowledge in a straightforward and linguistic way [55]. Due to its linguistic features, FZ has attracted considerable interests in the literature. Similar to the fault tree and expert systems, fuzzy logics adopt the *if-then* reasoning rule, which is a common and straightforward form of human knowledge. However, FZ stands out at its definition feature of using membership functions to describe the uncertainties and possibilities of events and rules [55]. As a result, FZ is able to easily incorporate uncertainties and possibilities, which are universal in data observation and decision making, into the diagnosis system. For example, a nonlinear fuzzy model [1] with transparent inner structure was used for the generation of six different symptoms in electro-pneumatic valve.

Due to the linguistic representation of human knowledge, FZ has shown its success in FDD [50]. A FZ system was developed in [80] for space monitoring and fault detection supported by the European Space Agency (ESA). In [76], a fuzzy spectral and spatial classifier was used for feature extraction. Fuzzy FDD was applied to induction motors, where the fuzzy bases were extracted from the current analysis of the fault modes [103]. In [65] a fuzzy-based classifier was developed to estimate types of actuator failure in aircraft, and a genetic algorithm was adopted to achieve an optimal fuzzy rule set for the classifier.

3) Principal Component Analysis (PCA) and Partial Least Squares (PLS): PCA and PLS are two typical multivariate statistical approaches in FDD [66]. Successful applications have been extensively reported in the literature. The first attempt of applying PCA in FDD can be found in [27], where overviews of using PCA and PLS in FDD were given. This method was extended to multiway PCA [74]. In order to handle nonlinearity in batch processes, a nonlinear PCA method was proposed in [25]. An integral statistical methodology combining PCA and discriminate analysis techniques was developed in [78]. In [27], PCA was discussed from a geometric point of view, and a methodology that analyzed fault subspace for process and sensor fault detection was addressed.

A major limitation of conventional PCA monitoring is that the PCA model is time-invariant, while most real processes are time-varying. Hence, the PCA model should also be recursively updated [101]. An adaptive monitoring approach using recursive PLS was presented in [94].

4) Other Supervised Methods: Other supervised methods include support vector machine (SVM), Bayesian classifier, and rough set. Recently, there have been a great deal of papers showing the application of SVM to FDD [70], including diagnosis of the bars in the machine [32]. In [60], a single-class SVM was developed for fault detection. In [46], the Bayes decision theory and Bayes minimum error classifier were applied to FDD. In [23], a two-step fuzzy/Bayesian formulation for changing point detection in time series was proposed and applied for incipient fault detection in dynamical systems. On decision tree analysis, a spatial decision tree was recently developed for movement monitoring [43]. A recent interesting study is the application of a hidden Markov model and parameter estimation techniques for condition monitoring of rotary machines [41].

C. Unsupervised Learning for FDD

The distinction between supervised and unsupervised learning is whether the training data provided for the "machine learner" has been labeled. Unsupervised learners are provided with the training data without classification tags. The unsupervised learner has to develop and select classification tags on its own.

Unsupervised algorithms usually seek out similarity between pieces of data in order to determine whether they can be characterized as forming a group (termed "cluster"). Thus, this process is also referred to as "clustering." In FDD, these different groups usually associate with different faults, and, ideally, each group is expected to have a one-to-one mapping to its own fault. However, the unsupervised algorithm does not guarantee this and may converge to solutions that are not optimal. For example, the selection of the number of clusters has been a potential difficult problem.

K-means is one of the best known and most popular clustering algorithms, which has found application to FDD [58]. Self-organizing neural networks such as ART network [7] have also been extensively used in fault diagnosis [96]. In addition, in [10] and [16], the integration of wavelets with ART networks was investigated for the development of diagnostic systems.

VI. Hybrid and Networked FDD in Industrial Automation

As these model-based, signal-based, and knowledge-based FDD techniques have their pros and cons, it is a trend that these three complementary techniques are usually integrated together to achieve a better performance. This is particularly true when the industrial processes have evolved from a set of loosely connected individual systems into a multitier networked automation system.

A. Multi-Tier FDD in Industrial Automation

In the fast changing industry automation, a large-scale complex automation system comprises three layers, and the data flows from bottom to top to drive different FDD algorithms. As illustrated in Fig. 5, these layers are given here.

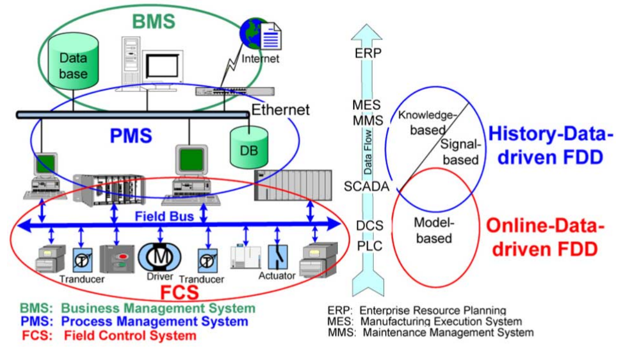

Fig. 5. Data flow and FDD in industrial automation.

1) Field control system (FCS): the field devices such as controllers, actuators and sensors are connected by correspondent field buses to form various control loops. Raw data is first sampled here and sent up for controlling and monitoring. Typical FCS are programmable logic controllers (PLC) and distributed control systems (DCS).
2) Process management system (PMS): the fundamental of this layer is a supervisory control and data acquisition (SCADA) system to collect and analyze the data distributed in FCSs. The safety and reliability are usually monitored at this layer, and appropriate supervisory control decisions and actions are taken to keep the process in a working state.
3) Business management system (BMS): the top layer usually consists of enterprise resource planning (ERP) system and maintenance management system.

A large-scale industrial system is a networked information system, where the raw data sampled at the lowest device level flows up to upper layers. Various data acquisition and processing tasks are carried out at different layers for different purposes. At the lowest FCS level, online data is processed in real time for model/signal-based FDD. At the middle PMS layer, a huge amount of online data are collected and stored over a longer period and processed later in a batch fashion. Depending on what type of data and how many data are available, the three FDD approaches reviewed in this paper are slotted into different layers but with quite a few overlaps.

B. Hybrid FDD

Different methods have their own advantages and disadvantages. The model-based FDD is able to detect and diagnose faults from small amount of online data in real time. Model-based methods have the ability to detect unknown type of fault, but it requires an explicit input-output model of the target system and its performance depends how good the model is. On the contrary, the signal-based and knowledge-based methods are supposed not to require an explicit or complete model of the system. Specifically, the signal-based methods focus on the analysis of the system's output signals with less attention to the dynamics of the input. Its performance may degrade when the system works in an unknown or unbalanced condition, while as the knowledge-based methods rely on the huge amount of high dimensional history data and are paid at the highest computational costs. As the knowledge-based FDD is on the basis of learn-by-example, its performance heavily relies on training data and is not good at detecting unknown faults [102].

It is commonly agreed that hybrid schemes would provide better solutions to a complex system. For instance, in model-based FDD, parameter identification is usually integrated into observer and parity space approaches to automate the process of modeling. In signal-based FDD, the time–frequency wavelet analysis is integrated with the MCSA in the frequency domain [5], [93]. In knowledge-based FDD, FZs are usually integrated into other methods. An ANFIS is a typical example [55], which sets up a neural network according to fuzzy rules, and the parameters of fuzzy rules are calibrated by backpropagation.

In particular, as fuzzy logics have easy representation of knowledge which usually is a drawback of other machine learning techniques, FZs are integrated into other methods. Statistical methods like PCA and PLS are also combined with NNs [44], [77], where PCA/PLS works as a feature extraction and selection tool to select statistical features and NN works as a classifier. Supervised and unsupervised methods can also be integrated. In [68] and [82], the unsupervised neural network with clustering was proposed. In [83], three techniques (i.e., PCA, FZ, and C-means clustering) are integrated to identify faults and develop operational strategy. The machine-learning techniques were also integrated into the qualitative methods. For example, a fuzzy expert system was proposed in [30].

Not only are various FDD techniques within the same category combined, but also there is a sign of integrating different methods crossover categories to overcome the cons of individual methods. In [53], various model-based, signal-based, and knowledge-based FDD are integrated into a distributed aero-engine health-monitoring system (DAME). In motor condition monitoring, the signal-based methods are integrated with model-based or knowledge-based methods, such as fuzzy logics [103] and neural networks [68]. In [84], combined with MCSA, fuzzy min-max (FMM) neural network and classification and regression tree (CART) were addressed to detect induction motor's faults. In [93], time–frequency analysis was used to extract the features of a rotary machine's vibration signal followed by a fuzzy sequential inference and diagnosis system to isolate the fault. The combination of model-based and signal-based FDD has shown its ability to detect faults under unbalanced conditions [36] and have attracted more attention recently.

C. FDD in Networked Control Systems (NCSS)

With the success of the real-time field bus network designed for control systems and the rapid development of communication networks, more non-real-time general networks, such as Ethernet and WiFi, are introduced into industrial automation, which opens up a new field of networked control system (NCS) [48]. In [53], a FDD system, DAME, was developed on grid computing that is a distributed data processing network. Recently, the emerging wireless sensor actuator networks for active flow control [6] and a recent WIDAGATE project [19] also witness this trend of NCSs. The wireless FDD also finds its promising application in building automation [62].

However, a most critical and important issue surrounding the increasing complexity in NCSs is to meet the requirements on system reliability. This makes networked FDD techniques receive more and more attention. It is known that the contention-based medium access control (MAC) and packet-exchange communication protocols widely accepted in NCSs introduce more uncertainties of delays and data losses into control loops and challenge the existing FDD. In networked FDD, much of attention has been paid to designing a fault-detection system robust to network-induced delays and packet losses [47]. A finite-state Markov chain is adopted to represent the dynamics of the network-induced delays and the control system is modeled as a Markov jumping system (MJS). Various FDD and optimization methods were proposed for MJS with the purpose to make FDD robustness to the network-induced delays, including Riccati equation methods [31] and linear matrix inequalities (LMIs) [47]. In [26], a knowledge-based fuzzy FDD was addressed for NCSs.

It is still an open question how a stochastic communication network affects the performance of NCSs and how a better FDD can be tailored for NCSs. As a disciplinary research area crossing control and communication, it is beneficial to bring the knowledge of communication networks (e.g., packet delay estimation and QoS metric) into FDD design, which could be a potential research direction in networked FDD. A pilot study is performed in [22], which made use of statistic features of MAC protocols to estimate the networked-induced delay and incorporated the delay information into the FDD design.

VII. Conclusion

In this paper, we have reviewed various analytic FDD methods from the perspective of how the data are processed. From a broad sense of information processing, all FDD systems are data-/signal-processing procedures with one search engine to check information redundancy between the data and explicit model or implicit knowledge. In this context, FDD methods are always data-driven. Depending on what kind of information (models, signals, or knowledge) are available and how the data and information redundancy are utilized, FDD methods are classified into three categories, namely model-based (online-data-driven) FDD, signal-based (data-driven) FDD, and knowledge-based (history-data-driven) FDD.

Given the extensive literature on the data-driven FDD, it is impossible to include all examples of them in a review due to space constraints. However, this paper sheds light on how the different methods relate and differ from one another within the unified framework of data processing. The trend of FDD in multitier industrial automation is also analyzed, and the potential research directions, such as hybrid methods and FDD in networked control systems, are presented.

Acknowledgment

Z. Gao, the corresponding author, would like to thank the anonymous reviewers, the associate editor, and the Editor-in-Chief for the constructive comments and ESAM, Northumbria University, for support.

REFERENCES

[1] P. Balle and D. Fuessel, "Closed loop fault diagnosis based on a no-linear process model and automatic fuzzy rule generation," *Eng. Applic. Artif. Intell.*, vol. 13, pp. 695–704, 2000.

[2] P. S. Barendse and P. Pillay, "A new algorithm for the detection of inter-turn stator faults in doubly-fed wind generators," in *Proc. 32nd Annu. Conf. IEEE Ind. Electron.*, Nov. 2006, pp. 1235–1240.

[3] M. Basseville, M. Abdelghani, and A. Benveniste, "Subspace-based fault detection algorithms for vibration monitoring," *Automatica*, vol. 36, no. 1, pp. 101–109, 2000.

[4] C. Bianchini, E. Fornasiero, T. N. Matzen, N. Bianchi, and A. Bellini, "Stator fault detection for multi-phase machines with multiple reference frames transformation," in *Proc. 35th Annu. Conf. IEEE Ind. Electron.*, 2009, pp. 3467–3470.

[5] A. Bouzida, O. Touhami, R. Ibtiouen, A. Belouchrani, M. Fadel, and A. Rezzoug, "Fault diagnosis in industrial induction machines through discrete wavelet transform," *IEEE Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4385–4395, Sep. 2011.

[6] K. Bur, P. Omiyi, and Y. Yang, "Wireless sensor and actuator networks: Enabling the nervous system of the active aircraft," *IEEE Commun. Mag.*, vol. 48, no. 7, pp. 118–125, Jul. 2010.

[7] C. A. Carpenter and S. Grossberg, "The art of adaptive pattern recognition by a self-organizing neural network," *Computer*, vol. 21, no. 3, pp. 77–88, 1989.

[8] T. Chai, Y. Zhang, H. Wang, C. Su, and J. Sun, "Data-based virtual unmodeled dynamics driven multivariable nonlinear adaptive switching control," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2154–2172, Dec. 2011.

[9] J. Chebil, G. Noel, M. Mesbah, and M. Deriche, "Wavelet decomposition for the detection and diagnosis of faults in rolling element bearings," *Jordan J. Mech. Ind. Eng.*, vol. 3, no. 4, pp. 260–267, 2009.

[10] B. H. Chen, X. Z. Wang, S. H. Yang, and C. Mcgreavy, "Application of wavelets and neural networks to diagnostic system development. 1. Feature extraction," *Comput. Chem. Eng.*, vol. 23, no. 7, pp. 899–906, 1999.

[11] C. Chen, B. Zhang, G. Vachtsevanos, and M. Orchard, "Machine condition prediction based on adaptive neuro-fuzzy and high-order particle filtering," *IEEE Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4353–4364, Sep. 2011.

[12] J. Chen, R. J. Patton, and G. P. Liu, "Optimal residual design for fault diagnosis using multi-objective optimization and genetic algorithms," *Int. J. Syst. Sci.*, vol. 27, no. 6, pp. 567–576, 1996.

[13] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Boston, MA, USA: Kluwer Academic, 1999.

[14] E. Y. Chow and A. S. Willsky, "Analytical redundancy and the design of robust failure-detection systems," *IEEE Trans. Autom. Control*, vol. AC-29, no. 7, pp. 603–614, Jul. 1984.

[15] T. W. S. Chow and S. Hai, "Induction machine fault diagnostic analysis with wavelet technique," *IEEE Trans. Ind. Electron.*, vol. 51, no. 3, pp. 558–565, Mar. 2004.

[16] S. M. A. Cruz and A. J. M. Cardoso, "Multiple reference frames theory: A new method for the diagnosis of stator faults in three-phase induction motors," *IEEE Trans. Energy Conv.*, vol. 20, no. 3, pp. 611–619, Sep. 2005.

[17] A. M. da Silva, R. J. Povinelli, and N. A. O. Demerdash, "Induction machine broken bar and stator short-circuit fault diagnostics based on three-phase stator current envelopes," *IEEE Trans. Ind. Electron.*, vol. 55, no. 3, pp. 1310–1318, Mar. 2008.

[18] X. Dai, T. Breikin, and H. Wang, "Modeling a complex aero-engine using reduced order models," in *Efficient Modeling and Control of Large-Scale Systems*. Berlin, Germany: Springer, 2010.

[19] X. Dai *et al.*, "Wireless communication networks for gas turbine engine testing," *Int. J. Distrib. Sensor Networks*, pp. 1–18, 2012.

[20] X. Dai, Z. Gao, T. Breikin, and H. Wang, "Disturbance attenuation in fault detection of gas turbine engines: A discrete robust observer design," *IEEE Trans. Syst., Man, Cybern. C*, vol. 39, no. 2, pp. 234–239, Mar. 2009.

[21] X. Dai, Z. Gao, T. Breikin, and H. Wang, "Zero assignment for robust H_2/H_∞ fault detection filter design," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1363–1372, Apr. 2009.

[22] X. Dai, Y. Yang, Y. Zhang, Z. Gao, and H. Wang, "Fault detection with MAC delay compensation in wireless sensor actuator networks," in *Proc. UKACC Int. Conf. Control*, Manchester, U.K., Sep. 2012, pp. 205–210.

[23] M. F. S. D'Angelo, R. M. Palhares, R. H. C. Takahashi, and R. H. Loschi, "Fuzzy/Bayesian change point detection approach to incipient fault detection," *IET Control Theory Applic.*, vol. 5, no. 4, pp. 539–551, 2011.

[24] V. T. Do and U. P. Chong, "Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain," *J. Mech. Eng.*, vol. 27, no. 9, pp. 655–666, 2011.

[25] D. Dong and T. J. McAvoy, "Batch tracking via nonlinear principal component analysis," *AIChE J.*, vol. 42, no. 8, pp. 2199–2208, 1996.

[26] H. Dong, Z. Wang, J. Lam, and H. Gao, "Fuzzy-model-based robust fault detection with stochastic mixed time delays and successive packet dropouts," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 365–376, Apr. 2012.

[27] R. Dunia and S. J. Qin, "Joint diagnosis of process and sensor faults using principal component analysis," *Control Eng. Practice*, vol. 6, pp. 457–469, 1998.

[28] S. Durovic, "Analysis of double-fed induction generator under electrical fault conditions," Ph.D. dissertation, Sch. Electr. Electron. eng., University of Manchester, Manchester, U.K., 2007.

[29] C. Edwards, S. K. Spurgeon, and R. J. Patton, "Sliding mode observers for fault detection," *Automatica*, vol. 36, pp. 541–553, 2000.

[30] S. M. El-Shal and A. S. Morris, "A fuzzy expert system for fault detection in statistical process control of industrial processes," *IEEE Trans. Syst., Man, Cybern. C, Applic. Rev.*, vol. 30, no. 2, pp. 281–289, Mar. 2000.

[31] H. Fang, H. Ye, and M. Zhong, "Fault diagnosis of networked control systems," *Annu. Rev. Control*, vol. 31, no. 1, pp. 55–68, 2007.

[32] R. Fang, "Induction machine rotor diagnosis using support vector machines and rough set," *Lect .Notes Comput. Sci.*, vol. 4114, pp. 631–636, 2006.

[33] P. J. Fleming and R. C. Purshouse, "Evolutionary algorithms in control systems engineering: A survey," *Control Eng. Practice*, vol. 10, no. 11, pp. 1223–1241, 2002.

[34] P. M. Frank and X. Ding, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *J. Process Control*, vol. 7, no. 6, pp. 403–424, 1997.

[35] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.

[36] A. Gandhi, T. Corrigan, and L. Parsa, "Recent advances in modeling and online detection of stator interturn faults in electrical motors," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1564–1575, May 2011.

[37] Z. Gao, T. Breikin, and H. Wang, "Reliable observer-based control against sensor failures for systems with time delays in both state and input," *IEEE Trans. Syst., Man Cybern. A*, vol. 38, no. 5, pp. 1018–1029, Sep. 2008.

[38] Z. Gao, T. Breikin, and H. Wang, "Discrete-time proportional and integral observer and observer-based controller for systems both with unknown input and output disturbances," *Optimal Control Applic. Methods*, vol. 29, no. 3, pp. 171–189, 2008.

[39] Z. W. Gao and H. Wang, "Descriptor observer approaches for multivariable systems with measurement noises and applications in fault detection and diagnosis," *Syst. Control Lett.*, vol. 55, pp. 304–313, 2006.

[40] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Trans. Ind. Inf.*, vol. 4, no. 4, pp. 271–279, Nov. 2008.

[41] O. Gerau, J. X. Xu, J. H. Zhou, and X. Li, "A physically segmented hidden Markov model approach for continuous tool condition monitoring: Diagnostics and prognostics," *IEEE Trans. Ind. Inf.*, vol. 8, no. 4, pp. 964–973, Nov. 2012.

[42] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Syst. Mag.*, vol. 8, no. 6, pp. 3–11, 1988.

[43] H. Ghasemzadeh and R. Jafari, "Physical movement monitoring using body sensor networks: A phonological approach to construct spatial decision trees," *IEEE Trans. Ind. Inf.*, vol. 7, no. 1, pp. 66–77, Feb. 2011.

[44] V. N. Ghate and S. V. Dudul, "Cascade neural-network-based fault classifier for three-phase induction motor," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1555–1563, May 2011.

[45] A. Y. Goharrizi and N. Sepehri, "A wavelet-based approach to internal seal damage diagnosis in hydraulic actuators," *IEEE Trans. Ind. Electron.*, vol. 57, no. 5, pp. 1755–1763, May 2010.

[46] M. Haji and H. A. Toliyat, "Pattern recognition-a technique for induction machines rotor broken bar detection," *IEEE Trans. Energy Conv.*, vol. 16, no. 4, pp. 312–317, Dec. 2001.

[47] X. He, Z. Wang, and D. Zhou, "Robust fault detection for networked systems with communication delay and data missing," *Automatica*, vol. 45, pp. 2634–2639, 2009.

[48] X. He, Z. Wang, Y. D. Ji, and D. H. Zhou, "Robust fault detection for networked systems with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 166–177, Jan. 2011.

[49] J. C. Hoskins, K. M. Kaliyur, and D. M. Himmelblau, "Fault diagnosis in complex chemical plants using artificial neural networks," *AIChE J.*, vol. 37, pp. 137–141, 1991.

[50] R. Isermann, "On fuzzy logic applications for automatic control, supervision and fault diagnosis," *IEEE Trans. Syst., Man Cybern. A, Syst. Humans*, vol. 28, no. 2, pp. 221–235, Jun. 1998.

[51] R. Isermann, "Process fault detection based on modeling and estimation methods—A survey," *Automatica*, vol. 30, pp. 397–404, 1984.

[52] R. Isermann, *Fault-Diagnosis Systems: An Introduction From Fault Detection to Fault Tolerance*. Berlin, Germany: Springer, 2006.

[53] T. Jackson, J. Austin, M. Fletcher, M. Jessop, B. Liang, A. Pasley, M. Ong, X. Ren, G. Allan, V. Kadirkamanathan, H. A. Thompson, and P. J. Fleming, "Distributed health monitoring for aero-engines on the GRID: DAME," in *Proc. IEEE Aerosp. Conf.*, Mar. 2005, pp. 3738–3747.

[54] I. M. Jaimoukha, Z. Li, and V. Papakos, "A matrix factorization solution to the H_-/H_∞ fault detection problem," *Automatica*, vol. 42, pp. 1907–1912, 2006.

[55] J.-S. R. Jang, C.-T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Englewood Cliffs, NJ, USA: Prentice-Hall PTR, 1997.

[56] P. Janik and T. Lobos, "Automated classification of power-quality disturbances using SVM and RBF networks," *IEEE Trans. Power Del.*, vol. 21, no. 3, pp. 1663–1669, Jul. 2006.

[57] S.-N. Ju, C.-L. Chen, and C.-T. Chang, "Constructing fault trees for advanced process control systems application to cascade control loops," *IEEE Trans. Reliabil.*, vol. 53, no. 1, pp. 43–60, Jan. 2004.

[58] K. Jyoti and S. Singh, "Data clustering approach to industrial process. Monitoring, fault detection and isolation," *Int. J. Comput. Applic.*, vol. 17, no. 2, pp. 41–45, 2011.

[59] S. N. Kavuri and V. Venkatasubramanian, "Neural network decomposition strategies for large scale fault diagnosis," *Int. J. Control*, vol. 59, no. 3, pp. 767–792, 1994.

[60] J. Kurek and S. Osowsk, "Support vector machine for fault diagnosis of the broken rotor bars of squirrel-cage induction motor," *Neural Computing Applic.*, vol. 19, pp. 557–564, 2010.

[61] S.-H. Liao, "Expert system methodologies and applications: A decade review from 1995 to 2004," *Expert Syst. With Applic.*, vol. 28, no. 1, pp. 93–103, 2005.

[62] Y. Liao, X. Dai, G. Liu, Y. Yang, and S. Yang, "Fault detection and eigenstructure optimization in IEEE 802.11 wireless sensor actuator networks for building automation," *Int. J. Adaptive Control Signal Process.* 2012, DOI: 10.1002/acs.2359.

[63] Q. Liu, T. Chai, H. Wang, and S. J. Qin, "Data-based hybrid tension estimation and fault diagnosis of cold rolling continuous annealing processes," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2284–2295, Dec. 2011.

[64] L. Ljung, *System Identification: Theory for the User*, 2nd ed. London, U.K.: Prentice-Hall, 1999.

[65] C. H. Lo, E. H. K. Fung, and Y. K. Wong, "Intelligent automatic fault detection for actuator failures in aircraft," *IEEE Trans. Ind. Inf.*, vol. 5, no. 1, pp. 50–55, Feb. 2009.

[66] M. D. Ma, D. S.-H. Wong, S. S. Jang, and S. T. Tseng, "Fault detection based on statistical multivariate analysis and microarray visualization," *IEEE Trans. Ind. Inf.*, vol. 6, no. 1, pp. 18–24, Feb. 2010.

[67] Y. Maki and K. A. Loparo, "A neural-network approach to fault detection and diagnosis in industrial processes," *IEEE Trans. Control Syst. Technol.*, vol. 5, no. 6, pp. 529–541, Dec. 1997.

[68] J. F. Martins, V. F. Pires, and A. J. Pires, "Unsupervised neural-network-based algorithm for an on-line diagnosis of three-phase induction motor stator fault," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, pp. 259–264, Jan. 2007.

[69] M. A. Massoumnia, "A geometric approach to the synthesis of failure detection filters," *IEEE Trans. Autom. Control*, vol. AC-31, no. 9, pp. 839–846, Sep. 1986.

[70] D. Meyer, F. Leisch, and K. Hornik, "The support vector machine under test," *Neurocomputing*, vol. 55, pp. 169–186, 2003.

[71] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors: A review," *IEEE Trans. Energy Conv.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[72] K. G. Narendra, V. K. Sood, K. Khorasani, and R. Patel, "Application of a radial basis function (RBF) neural network for fault diagnosis in a HVDC system," *IEEE Trans. Power Syst.*, vol. 13, no. 1, pp. 177–183, Jan. 1998.

[73] K. Niida, "Expert system experiments in processing engineering," in *Proc. Inst. Chem. Eng. Symp. Series*, 1985, pp. 529–583.

[74] P. Nomikos and J. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE J.*, vol. 40, no. 8, pp. 1361–1375, 1994.

[75] R. J. Patton, P. Frank, and R. Clark, *Fault Diagnosis in Dynamic Systems: Theory and Application*. London, U.K.: Prentice-Hall, 1989.

[76] A. Picon, O. Ghita, P. F. Whelan, and P. M. Iriondo, "Fuzzy spectral and spatial feature integration for classification of nonferrous materials in hyperspectral data," *IEEE Trans. Ind. Inf.*, vol. 5, no. 4, pp. 483–494, Nov. 2009.

[77] S. J. Qin and T. J. MacAvoy, "Nonlinear PLS modeling using neural networks," *Comput. Chem. Eng.*, vol. 16, pp. 379–391, 1992.

[78] A. Raich and A. Cinar, "Diagnosis of process disturbances by statistical distance and angle measures," *Comput. Chem. Eng.*, vol. 21, pp. 661–673, 1997.

[79] S. Rajagopalan, J. A. Restrepo, J. M. Aller, T. G. Habetler, and R. G. Harley, "Non stationary motor fault detection using recent quadratic time-frequency representations," *IEEE Trans. Ind. Applic.*, vol. 44, no. 3, pp. 735–744, May–Jun. 2008.

[80] R. A. Ribeiro, "Fuzzy space monitoring and fault detection applications," *J. Decision Syst.*, vol. 15, pp. 267–286, 2006.

[81] J. Ribrant and L. M. Bertling, "Survey of failures in wind power systems with focus on Swedish wind power plants during 1997-2005," *IEEE Trans. Energy Conv.*, vol. 22, no. 1, pp. 167–173, Mar. 2007.

[82] R. R. Schoen, B. K. Lin, T. G. Habetler, J. H. Schlag, and S. Farag, "An unsupervised, on-line system for induction motor fault detection using stator current monitoring," *IEEE Trans. Ind. Applic.*, vol. 31, no. 6, pp. 1280–1286, 1995.

[83] Y. M. Sebzalli and X. Z. Wang, "Knowledge discovery from process operational data using PCA and fuzzy clustering," *Eng. Applic. Artif. Intell.*, vol. 14, pp. 607–616, 2001.

[84] M. Seera, C. P. Lim, D. Ishak, and H. Singh, "Fault detection and diagnosis of induction motors using motor current signature analysis and a hybrid FMM-CART model," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 23, no. 1, pp. 97–108, Feb. 2012.

[85] B. Shafai, C. T. Pi, and S. Nork, "Simultaneous disturbance attenuation and fault detection using proportional integral observers," in *Proc. Amer. Control Conf.*, 2002, vol. 2, pp. 1647–1649.

[86] J. Sottile, Jr and L. E. Holloway, "An overview of fault monitoring and diagnosis in mining equipment," *IEEE Trans. Ind. Applic.*, vol. 30, no. 5, pp. 1326–1332, Sep.–Oct. 1994.

[87] A. Srividya, A. K. Verma, and B. Sreejith, "Automated diagnosis of rolling element bearing defects using time-domain features and neural networks," *Int. J. Mining, Reclamation Environ.*, vol. 23, no. 3, pp. 206–215, 2009.

[88] C. Tan and C. Edwards, "Sliding mode observers for detection and reconstruction of sensor faults," *Automatica*, vol. 38, pp. 1815–1821, 2002.

[89] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "A review of process fault detection and diagnosis part I: Quantitative model-based methods," *Comput. Chem. Eng.*, vol. 27, no. 3, pp. 293–311, 2003.

[90] D. Volante, "Condition monitoring for rotational machinery," M.Sc. thesis, Dept. Computing and Software, McMaster Univ., Hamilton, ON, Canada, 2011.

[91] H. Wang, T. Y. Chai, J. Ding, and M. Brown, "Data driven fault diagnosis and tolerant control," *Automatica Sinica*, vol. 35, no. 6, pp. 739–747, 2009.

[92] H. Wang and S. Daley, "Actuator fault diagnosis: An adaptive observer-based technique," *IEEE Trans. Autom. Control*, vol. 41, no. 7, pp. 1073–1078, Jul. 1996.

[93] H. Wang and P. Chen, "Fuzzy diagnosis method for rotating machinery in variable rotating speed," *IEEE Sensors J.*, vol. 11, no. 1, pp. 23–34, Jan. 2011.

[94] X. Wang, U. Kruger, and G. W. Irwin, "Process fault diagnosis using recursive multivariate statistical process control," in *Proc. 16th IFAC World Congress*, 2005, pp. 1849–1895.

[95] K. Watanabe and D. M. Himmelblau, "Instrument fault detection in systems with uncertainties," *Int. J. Syst. Sci.*, vol. 13, no. 2, pp. 137–158, 1982.

[96] J. R. Whiteley and J. F. Davis, “A similarity-based approach to interpretation of sensor data using adaptive resonance theory,” *Comput. Chem. Eng.*, vol. 18, no. 7, pp. 637–661, 1994.

[97] M. Witczak, *Modelling and Estimation Strategies for Fault Diagnosis of Non-Linear Systems: From Analytical to Soft Computing Approaches*, 1st ed. Berlin, Germany: Springer, 2007.

[98] W. Yang, P. J. Tavner, C. J. Crabtree, and M. Wilkinson, “Cost-effective condition monitoring for wind turbines,” *IEEE Trans. Ind. Electron.*, vol. 57, no. 1, pp. 263–271, Jan. 2010.

[99] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, “Inchoate fault detection framework: Adaptive selection of wavelet nodes and cumulant orders,” *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 685–695, Jun. 2012.

[100] J. Zhang and P. Roberts, “Process fault diagnosis with diagnostic rules based on structural decomposition,” *J. Process Control*, vol. 1, no. 5, pp. 259–269, 1991.

[101] Y. Zhang, H. Zhou, S. J. Qin, and T. Chai, “Decentralized fault diagnosis of large-scale processes using multiblock Kernel partial least squares,” *IEEE Trans. Ind. Inf.*, vol. 6, no. 1, pp. 3–10, Feb. 2010.

[102] J. H. Zhou, D. H. Zhang, C. H. Ooi, M. Luo, S. Mao, and D. W. Wang, “Comparative study of data driven and model based approaches of rotary machines fault detection and diagnosis,” *SIMTech Tech. Rep.*, vol. 11, no. 4, pp. 195–201, 2010.

[103] F. Zidani, D. Diallo, M. E. H. Benbouzid, and R. Nait-Said, “A fuzzy-based approach for the diagnosis of fault modes in a voltage-fed PWM inverter induction motor drive,” *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 586–593, Feb. 2008.

Xuewu Dai (M’11) received the B.Eng. degree in electronic engineering and M.Sc. degree in computer science from Southwest University, Chongqing, China, in 1999 and 2003, respectively, and the Ph.D. degree from The University of Manchester, Manchester, U.K., in 2008. His doctoral work was on observer-based parameter estimation and condition monitoring for gas turbine engines.

In 2002, he joined the School of Electronic and Information Engineering, Southwest University, and received a lectureship in 2003. From 2009 to 2012, he held research positions with the University of College, London, U.K., and the University of Oxford, Oxford, U.K., respectively. His main research interests include robust adaptive signal processing, state estimation and its applications to condition monitoring of aero-engines and wind generators, wireless communication, wireless sensor networks, and wireless networked control systems.

Zhiwei Gao (SM’08) received the B.Eng. degree in electric engineering and automation and M.Eng. and Ph.D. degrees in systems engineering from Tianjin University, Tianjin, China, in 1987, 1993, and 1996, respectively.

From 1987 to 1990, he was with Tianjin Electric Drive and Design Institute as an Assistant Engineer working on industrial projects. From 1996 to 1998, he was with the Department of Mathematics, Nankai University, as a Postdoctoral Researcher. In 1998, he joined the School of Electric Engineering and Automation and received a professorship in control science and engineering in 2001. Before joining the Faculty of Engineering and Environment at Northumbria University in 2011, he held lecturing and research positions with the City University of Hong Kong, University of Manchester Institute of Science and Technology, University of Duisburg-Essen, University of Manchester, University of Leicester, University of Liverpool, and Newcastle University. His research interests include data-driven modelling, estimation and filtering, fault diagnosis, fault-tolerant control, intelligent optimisation, large-scale systems, singular systems, distribution estimation and control, renewable energy systems, gene regulation network and biomedical systems.

Prof. Gao organized three special issues on complex systems for *Open Automation and Control Systems Journal, Journal of Applied Mathematics* and the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS as a lead Guest Editor. He is the editor-in-chief of the *International Journal of Systems, Signal, Control and Engineering Applications*. He is presently an associate editor for the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY. He received the Tianjin Natural Scientific Prize in 2000 and the Alexander von Humboldt Research Fellowship in 2004.