Multivariate Anomaly Detection in Real-World Industrial Systems

Xiao Hu, Raj Subbu, Piero Bonissone, Hai Qiu, Naresh Iyer

Abstract—Anomaly detection is a critical capability enabling condition-based maintenance (CBM) in complex real-world industrial systems. It involves monitoring changes to system state to detect "anomalous" behavior. Timely and reliable detection of anomalies that indicate faulty conditions can help in early fault diagnostics. This will allow for timely maintenance actions to be taken before the fault progresses and before it causes secondary damage to the system leading to downtime. When an anomaly is identified, it is important to isolate the source of the fault so that appropriate maintenance actions can be taken. In this paper, we introduce effective multivariate anomaly detection techniques and methods that allow fault isolation. We present experimental results from the application of these techniques to a high-bypass commercial aircraft engine.

I. INTRODUCTION

In recent years there has been a growing interest and investment in condition-based maintenance (CBM) in aerospace, manufacture, transportation and other complex industrial systems. Different from the traditional time-based preventive maintenance or failure-triggered corrective maintenance, CBM has been defined as a set of maintenance processes and capabilities derived from assessment of system condition, obtained from in-situ, non-invasive test, operating and condition measurements [1]. Minimizing total cost of ownership, increasing system/asset availability and minimizing system downtime while increasing safety are some of the key benefits of implementing CBM. The deployment of CBM strategies requires two essential capabilities—diagnostics and prognostics. The focus of diagnostics is timely and accurate detection of anomalies in system condition, and identification of the root causes responsible for the anomaly. Most diagnostic methods rely on discernable changes in observable parameters in order to detect faults. Therefore, anomaly detection is a critical capability to enable CBM in complex industrial systems. Early and reliable detection of anomalies that indicate faulty conditions can help in early fault diagnostics. This will allow for timely maintenance actions to be taken before the fault progresses and before it causes secondary damage to the system leading to system downtime. The focus of prognostics is the prediction of remaining useful life until the next maintenance event, given a current faulty or deteriorated system state.

From a modeling perspective, there are anomaly detection methods that require accurate system models [2]-[5], which often demand fundamental understanding of the physical system. This is sometimes difficult to acquire. These model-based approaches identify anomalies by determining if the relationships among variables violate basic rules of physics, or by analyzing the residuals between the actual measurements of system variables and the estimation of system variables by the model. On the other hand, there are methods that do not assume any a priori domain knowledge of the system and rely only on historic data, although any domain knowledge can be leveraged when they are available. These empirical model and data-driven approaches tend to detect anomalies by learning the envelope of normality from historic data. Empirical approaches are applicable generally across platforms, and the focus of anomaly detection in CBM, and consequently of this paper. From this point on anomaly detection in the rest of paper refers to anomaly detection using empirical approaches.

Depending on the number of variables considered, anomaly detection techniques can be categorized into univariate and multivariate. There have been a number techniques suggested in the literature for detecting anomalies in one-dimensional time-series data in monitored systems [6]-[9]. However, these univariate techniques cannot capture the interaction and inter-relationship among the various system variables being monitored as a whole. Often times, the change of system state can be more effectively identified through monitoring of the changes in covariance among sensors/variables. Multivariate techniques can be applied in such situations. Research in multivariate anomaly detection has been very active in the last decade. Roughly, they can be organized into two categories—statistical approaches [10] and neural network based approaches [11]. The statistical approaches include probabilistic/Gaussian Mixture Modeling (GMM), Hidden Markov Models, hypothesis testing and clustering techniques. Among neural network based approaches, principally Support Vector Machines (SVM), Adaptive Resonance Theory (ART) approaches, and Self-Organizing Maps (SOM) approaches. In this paper, we will describe statistics based anomaly detection techniques [12], which are very promising in identifying anomalies from multi-dimensional data of complex industrial system as well as implicating responsible variables that cause the anomaly.
The remainder of the paper is organized as follows. Section II briefly defines the problem of anomaly detection in industrial systems. Section III introduces the multivariate anomaly detection techniques, Hotelling’s T² statistic and its extension. Section III presents experimental results on data from a high bypass commercial aircraft engine. We then discuss the constraints and potential improvement to the techniques in Section IV. Section V concludes the paper.

II. PROBLEM DEFINITION
We first address the problem definition of anomaly detection in industrial systems. Fig. 1 generally depicts the components of a typical system. The figure shows a controlled system and indicates the different sources of failures in it. In this system, the control signal \( u_c \) from controller is converted into a physical control input/signal \( u_p \) by actuator. The actual system output \( y_p \), some of the system state variables \( x_p \) and \( u_p \) are sensed as \( x_s \), \( y_s \) and \( u_s \) by a set of sensors. The state estimator predicts the remainder of the state variables of interest \( x_e \). \( x_s \), \( y_s \) and \( u_s \) are the data sources available to the empirical anomaly detection approaches. Any failure or malfunction that occurs at any component may cause the change of \( x_s \), \( y_s \) or \( u_s \), indicating anomalies. The challenge of fault diagnosis is to detect anomalies from the various sensor measurements, identify which variables are responsible for causing the anomalous sensed system behavior, and then isolate the failing or failed component so that appropriate maintenance actions can be taken.

![Figure 1. A general controlled system.](image)

III. HOTELLING’S \( T^2 \) STATISTIC AND ITS VARIANT

A. Traditional Hotelling’s \( T^2 \) Statistic
The Hotelling’s \( T^2 \) statistic, alternatively T2, is a well-known multivariate statistical technique in the process quality control community. First proposed by Harold Hotelling [13], it is a generalization of Student’s \( t \) statistic that is used in multivariate hypothesis testing. It is defined with respect to a set of \( p \) variables \( x = (x_1, x_2, ..., x_p) \) having mean values \( \mu = (\mu_1, \mu_2, ..., \mu_p) \) and \( p \times p \) covariance matrix, \( W \). Hotelling’s T2 statistic is given as:

\[
T^2 = (x - \mu) W^{-1} (x - \mu)
\]

where \( W = \sum_{i=1}^{n} (x_i - \mu_i)(x_i - \mu_i)'/ (n-1) \) is formed from \( n \) snapshots taken of the system-associated variables [14]. So, in a typical on-line monitoring process, sensor measurements are collected sequentially and a decision of when a change occurs is made based on the value of \( T^2 \).

Expected behavior of the Hotelling’s T2 statistic is based on the assumption that \( x \) constitutes a joint \( p \)-variate Gaussian distribution. In practice one never sees such a distribution. However, the Hotelling’s T2 statistic and many other statistics based on the same assumption may contribute immense value to system analysis.

When using the Hotelling’s T2 statistic on system monitoring, it is generally advisable to provide a training period with historic data taken from sensors observing the system-associated variables when the system is reasonably believed to be in stable operation. During such an interval, the mean values \( \mu \) are presumed to exist and may be derived by averaging. It is a change in the \( p \)-dimensional covariance matrix \( W \), or in the mean values \( \mu \) that causes the one-dimensional Hotelling’s T2 statistic to exhibit a detectable change.

B. Variable Contribution
Once a change/anomaly in the operation of complex system has been identified by the Hotelling’s T2 statistic, we need to isolate the system-associated variables driving the departure from normality. One way to perform this step is to first express the Hotelling’s T2 statistic in terms of its principal components. This is done by first representing T2 as:

\[
\tau^2 = \sum_{a=1}^{n} \frac{t_a^2}{\lambda_a} = \sum_{a=1}^{n} \frac{s_a^2}{\lambda_a}
\]

where \( \lambda_a \), \( a = 1, 2, ..., n \), are the eigenvalues of covariance matrix \( W \), and \( t_a \) are the scores from the principal component transformation. \( s_a^2 \) is the variance of \( t_a \) (the variance of the principal components are the eigenvalues of \( W \)) [15]. Each score \( t_a \) can be expressed as

\[
t_a = p_a' (x - \mu) = \sum_{j=1}^{n} p_{a,j} (x_j - \mu_j)
\]

where \( p_a \) is the eigenvector of \( W \) corresponding to \( \lambda_a \), and \( p_{a,j} \), \( x_j \), \( \mu_j \) are elements of the corresponding vector, associated with the \( j^{th} \) variable. By this representation, the
contribution of each variable $x_j$ to the score of the principal component $a$ is found to be $p_{a,j}(x_j - \mu_j)$. This information can be used to identify the variable(s) that had the strongest impact to the anomaly measured by the Hotelling’s T2 metric. For each one of the normalized scores $(t_a / s_a)^2$ with high values we calculate the variable contribution, but keep only the contributions with the same sign as $t_a$. If $m \leq n$ high scores are identified, then one can calculate the contribution $cont_{a,j}$ of a variable appearing in the normalized score $(t_a / s_a)^2$. We have

$$cont_{a,j} = (t_a / s_a)^2 p_{a,j}(x_j - \mu_j).$$

$cont_{a,j}$ is set equal to zero if it is negative, which means that its sign is opposite to the value of the score $t_a$. We then calculate the total contribution of variable $x_j$ by summing

$$CONT_j = \sum_{a=1}^{n} (cont_{a,j}).$$

The Hotelling’s T2 statistic may now be written as a sum of contributions of all variables over all of the scores as

$$T^2 = \sum_{a=1}^{n} \sum_{j=1}^{m} (t_a / s_a)^2 p_{a,j}(x_j - \mu_j)$$

$$= \sum_{a=1}^{n} \sum_{j=1}^{m} \frac{t_a}{s_a} [p_{a,j}(x_j - \mu_j)]$$

As pointed out in [15], the key is to only sum up the contributions over the high scores and keeping the positive $cont_{a,j}$. It reveals the variables responsible for an increase both of the value of T2 and the value of high normalized score $(t_a / s_a)^2$.

When a large number of system-associated variables are monitored (it is not uncommon to have more than 50 monitored variables in a complex industrial system) and some of those variables are highly correlated, it becomes impractical to employ traditional T2 approaches. Those highly correlated variables may lead to ill-conditioned data, which cause T2 to become very sensitive to deviations of components corresponding to small eigenvalues. The inversion of the covariance matrix may also become a problem. To handle such multivariate problems in high dimensional data space, multivariate statistical approach based on principal component analysis and partial least squares or projection to latent structures have been developed [16]-[18]. They are capable of compressing data and reducing dimensionality so that essential information is retained, and they are also able to handle noise and correlation to extract true information effectively.

### IV. EXPERIMENT RESULTS

We illustrate the application of Hotelling’s T2 techniques for anomaly detection and differentiate between sensor fault and system fault on synthetic and real data from a high-bypass commercial aircraft engine. In this feasibility study, in order to more conveniently simulate sensor faults and system component faults, we leveraged a highly accurate Component Level Model (CLM). It is a physics-based thermodynamic model that is used to simulate the performance of a real-world aircraft engine. Flight conditions such as altitude, Mach number, ambient temperature, engine fan speed, and a large variety of model parameters, such as module efficiency and flow capacity are inputs to the CLM (see Fig. 2). The outputs of the CLM are the values for pressures, core speed and temperatures at various locations of engine. The outputs simulate sensor measurements. Realistic values of sensor noise can be added after the CLM calculation. In this study, a steady state CLM model for a commercial, high-bypass, turbofan engine is used. The objective is to use engine data collected under cruise conditions to monitor engine health changes.

Real engine monitoring snapshot data were collected from a turbofan engine over a period, at the end of which it was overhauled due to deterioration. Fig. 3(a) shows the normalized sensor measurements of the principal three of the eleven variables: exhaust gas temperature – EGT, fuel flow – WFM, and core speed – N2. From the figure, we can see that there is no obvious trend or shift in sensor measurements as the engine deteriorates. First, data from the first N flight cycles (points) were used as reference to estimate the sample mean $\mu$ and sample covariance matrix $W$. Next, the T2 statistic was computed directly for all the data based on (1), including EGT, WFM, N2 and some other key parameters. In Fig. 3(b), we can see the upward trending of the T2 score, showing higher values as we approach the engine deterioration level that caused its overhaul.

As an example of the technique to determine when the complex system departs from normality and to assess the contributions of the responsible variables, which help discriminate between sensor fault and system fault, consider that there are nine critical sensors $x_1, x_2, \ldots, x_9$ from the engine model that exhibit sequential values as graphed in Fig. 4(a). The data of the 9 variables between flight 1 and
200 were generated while the model was simulated under normal flight condition without loss of efficiency and flow. We can regard those data "no fault (normal)" data, which were used to build the reference model for T2 analysis. To simulate sensor fault, we intentionally modified some of the sensor readings at different times (flight number 130, 150, 170 and 185, respectively), as indicated by the red arrows for N2, PS, T3 and EGT. From flight 201, we reconfigured the model with certain degree of loss of flow and efficiency to simulate system component fault, shown in Fig. 4(a). The next step is to illustrate the contributions from the nine variables to the Hotelling's T2 statistic. This may be done in several ways. The contribution magnitudes and relative significance may be displayed as in Fig. 4(b) using color-coding for aiding the operator in interpreting the results. The color-coded 280 by 9 grid displays the relative contributions on the ordinate from the nine variables listed on the abscissa. T2 statistic is calculated based on the time series data from the 9 sensors, also shown in Fig. 4(b). The location of the four peaks between flight number 100 and 200 in the T2 statistic plots corresponds to where the sensor faults were injected. The sensor variables with high contribution are implicated for the change in T2 score. The sensors identified through variable contribution analysis are the same sensors with injected faults. Then, T2 statistic ramps up quickly since system fault starts from flight 200. The contribution analysis recognizes multiple sensors responsible for the T2 scores exceeding the predefined alerting threshold. Since the odds that multiple sensors have sensor fault simultaneously is so small that we can almost conclude based on heuristics that the anomalies in T2 scores, which were caused by multiple sensor variables identified through contribution analysis indicate system component fault or change of system operating condition. By knowing the operating condition of the system, this technique can be utilized to discriminate between a sensor fault and system fault, respectively.

V. DISCUSSION AND FUTURE WORK

A. Prerequisite of Deployment of Hotelling's T2 techniques

Both Hotelling's T2 statistic and its variants leverage covariance information. For them to work properly, there has to be dependencies (correlations or interactions) among the variables being monitored. This prerequisite is generally met for most of complex industrial systems, such as sensor data collected from turbine, aircraft engine, and locomotives. However, it is worthwhile to confirm the correlations of system-associated variable before applying any of these techniques.

B. Improvements in Computing the Covariance Matrix

To increase the Hotelling’s T2 score’s sensitivity in detecting a shift or drift in the mean vector, we can compute the covariance matrix estimator by constructing a successive differences estimator \[ \sum_{i=1}^{n} \left( V_{ij}^{*} - V_{ij} \right) \] for \( i = 1, 2, \ldots, n - 1 \) and is used to form

\[ S^{*} = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} V_{ij}^{*} V_{ij} . \]  

Another approach is to compute the Minimum Volume Ellipsoid (MVE) [20], as a way to provide an estimate for the mean vector and covariance matrix that is robust to outliers. Given a collection of observations, we compute the mean vector at the center of the minimum volume ellipsoid covering half of the observations, and the covariance matrix is determined by the same ellipsoid (multiplied by a correction factor to obtain consistency at multinomial distributions).
C. Improvements in Identifying Variable Contributions to the Hotelling’s T2 score

We could complement the variable implication with metrics related to Entropy and minority-decision cardinality, which could be used to quantify the confidence in the identification of the variable contributions. Examples of these metrics are:

\[
E_i = -\frac{1}{\ln(N)} \sum_{j=1}^{N} M_{a,j} \ln(M_{a,j}) \quad (8)
\]

\[
m_i = 1 - \max_j \left(M_{a,j}^{'}\right) \quad (9)
\]

where \(M_{a,j}^{'} = \frac{\rho_{a,j}(x_j - \mu_j)}{\sum_{j=1}^{N} M_{a,j}}\) is the magnitude (absolute value) of the variable contribution and \(M_{a,j} = M_{a,j} \sum_{j=1}^{N} M_{a,j}\) is the normalized magnitude.

As the example shown in Fig. 4(c), the variable contributions at flight 130 will have lower entropy \(E\) based on (8) than at flight 265, as the contribution from EGT dominates the others at flight 130, and the contribution from other variables are also relatively significant at flight 265.
VI. CONCLUSIONS

As one of the key enabling technologies for Condition-Based Maintenance, anomaly detection plays a critical role in facilitating diagnostics and prognostics to increase system/asset availability and minimize system downtime while increasing safety margins. In this paper, we presented multivariate anomaly detection techniques, which take into consideration covariance changes among multiple sensors/variables in condition monitoring of complex industrial systems. Experiments on models and data from a high bypass, turbofan aircraft engine model show promising results. The techniques introduced not only can detect anomalies from sensor data, but also can identify the variables responsible for the anomalies.

REFERENCES


