Kalpa Publications in Computing
Volume 4, 2018, Pages 21–36
28th International Workshop on
Principles of Diagnosis (DX'17)

A combined model-based and data-driven approach for monitoring smart buildings

Hamed Khorasgani and Gautam Biswas

Institute of Software Integrated Systems, Vanderbilt University, USA
{hamed.g.khorasgani,gautam.biswas}@vanderbilt.edu

Abstract

This paper combines a residual-based diagnosis approach and an unsupervised anomaly detection method to develop a hybrid methodology for monitoring smart buildings for which complete models are not available. The proposed method combines data mining approach and model-based diagnosis to update a diagnosis reference model and improve the overall diagnostics performance. To estimate the likelihood of each potential fault in complex systems like smart buildings, the dependencies between components and, therefore, the monitors should be considered. In this work, a tree augmented naive Bayesian learning algorithm (TAN) is used for the classification. We demonstrate and validate the proposed approach using a data-set from an outdoor air unit (OAU) system in the Lentz public health center in Nashville.

1 Introduction

A smart building is defined as the integration of building management, and telecommunication technologies [21]. The Internet of Things (IoT) [18] has brought a new generation of smart buildings with sensor networks that can monitor a number of different building variables, such as temperature, humidity, air flow, and power usage. This further enables us to detect and isolate faults when they occur, and develop control algorithms that are designed to reduce energy consumption, improve operating conditions and tenant satisfaction, and reduce unplanned downtime. To achieve these goals, early fault detection and isolation (FDI) is critical in any smart building monitoring strategy. Effective FDI supports prompt remedial actions and reduces unnecessary wastage of resources, decreases residents discomfort, and prevents extended downtime.

Traditional approaches to FDI typically rely on a model that defines nominal behavior of a system. Model-based approaches such as graphical methods [16, 5], observer-based FDI [1], and parity equations [8] have been successfully applied for FDI in dynamic systems. Provan [19] developed a diagnosis model for FDI in heating, ventilation and air conditioning (HVAC) systems in smart buildings. The model-based methods are computationally efficient. Moreover, it is easy to understand and interpret the diagnosis results when we apply these approaches. However, for complex systems, it is expensive and sometimes infeasible to generate an accurate model for the entire system. When the model is incomplete, data-driven monitoring approaches

[23] can be used as a promising supplement to model based fault detection and isolation. A data-driven monitoring algorithm uses the system measurements to observe the system behavior, recognizes unusual situations during its operation, and informs experts to update operations procedures.

Several researchers have combined model-based diagnosis with data-driven approaches to achieve better diagnosis performance. Mack et al [13] combined a diagnosis reference model developed by domain experts with a tree augmented naive Bayesian (TAN) learning algorithm to develop a framework that combines the expert knowledge and historical data to improve the accuracy in differentiating between nominal and faulty situations. Data has also been used to learn uncertain and incomplete models [22]. Jung et al [9] used a model-based approach for fault detection. To achieve better fault isolation, they used the residuals outputs in previous fault scenarios to generate models for different fault modes. For each fault mode, they trained a one-class support vector machine (1-SVM) to determine if a new sample (residual output) belongs to the fault mode or not. If the new sample did not belong to any of the fault modes, it was labeled as a likely unknown fault. The classifiers were expected to become more accurate as more data is obtained overtime.

Sheibat-Othman et al [20] used support vector machines (SVM) for fault detection and applied an observer-based diagnosis approach for fault isolation. When the SVM detects no fault in the system, they use the data to update the observer parameters. Narasimhan et al [17] used TRANSCEND diagnosis approach [16] to reduce the set of possible faults and then applied a data-driven approach to best distinguish among the remaining faults. In this paper, our goal is to develop a hybrid diagnosis approach that combines the use of historical data with the available physics-based knowledge of the system to achieve better diagnosis performance in smart buildings with incomplete models. By combining model-based diagnosis and data-driven anomaly detection, we can detect and isolate faults that was not possible with pure model-based diagnosis approaches because the models are incomplete.

The rest of the paper is organized as follows. Section 2 discusses our proposed combined methodology for fault detection and isolation in complex systems with incomplete models. Section 3 presents the case study using a subsystem of Lentz Public Health Center. Section 4 presents conclusions and discusses the future work.

2 Methodology

We have developed a multi-step diagnosis methods to distinguish outlier behaviors and link them to anomalies or faults in smart buildings. Figure 1 describes a three step approach for monitoring smart buildings: (1) using a model-based diagnosis approach to generate the initial diagnosis reference model; (2) unsupervised anomaly detection using a density-based clustering approach; (3) identifying the undetected outlier groups, deriving the significant features that characterize each outlier group from the nominal and using these significant features to complete the diagnosis reference model. We describe each step in greater detail below.

2.1 Model-based Diagnosis

Residuals represent redundancies in the system equations. The model-based diagnosis unit in in Figure 1, applies the set of generated residuals, $R = \{r_1, ..., r_l\}$, to detect and isolate system faults. These residuals are associated to the sections of a building that the analytical model is available and we can use analytic redundancy relationships to derive the residuals. To address

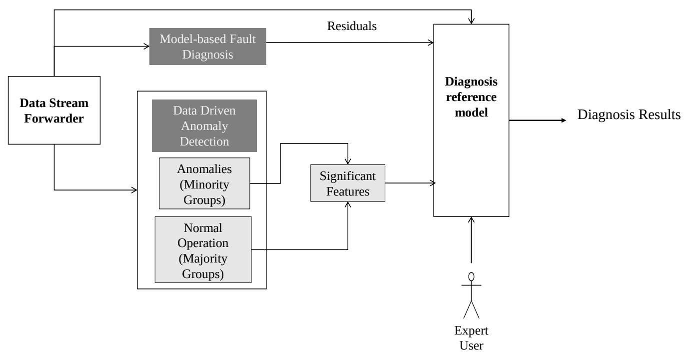

Figure 1: A combined model-based and data-driven approach for monitoring smart buildings.

noise and uncertainty and generate robust diagnosis results, we apply a Z-test [4] to determine when change in a residual value, $r \in R$, is statistically significant.

We consider the last N_2 residual values to compute the mean value of residual distribution (assumed to be a normal distribution):

$$\mu_r(k) = \frac{1}{N_2} \sum_{i=k-N_2+1}^{k} r(i). \tag{1}$$

The last N_1 samples (typically, $N_1 \gg N_2$) to compute the variance:

$$\sigma_r^2(k) = \frac{1}{k - N_1 - 1} \sum_{i=k-N_1+1}^{k} (r(i) - \mu_r(k))^2. \tag{2}$$

The confidence level for the Z-test, α, determines the bounds, z_-, and z_+, and, therefore, the sensitivity of the residuals.

$$P(z_- < (r(k) - \mu_r(k)) < z_+) = 1 - \alpha. \tag{3}$$

The Z-test is implemented as follows:

$$m_r = \begin{cases} 0 & if \qquad z_- < r(k) - \mu_r(k) < z_+ \\ 1 & otherwise. \end{cases} \tag{4}$$

where m_r represents the monitor derived from residual r. Ideally, a monitor fires when at least one fault has occurred in the system, and the set of monitors derived from system equations, $M_r = \{m_r | \forall r \in R\}$, are adequate to detect and isolate all the system faults. When the system model is incomplete and the set of monitors derived from system equations, M_r is not enough to achieve complete diagnosability performance, the system's historical data can be used to derive additional monitors.

2.2 Data-driven anomaly detection

2.2.1 Clustering the data objects:

Clustering is a process of partitioning a set of objects into clusters such that objects in the same cluster are more similar to each other than objects in different clusters according to some defined criteria [14]. Clustering methods can be applied to unsupervised anomaly detection. Li et al [12] used a density-based clustering approach to detect anomalous flights based on onboard-recorded flight data. In previous work [3] we used a hierarchical clustering to identify anomalies in a spacecraft telemetry data.

Typically, a large majority of the data-points will represent nominal operations of a building but a small subset may represent anomalous and faulty behaviors. For unsupervised learning in step 2, we use the density-based spatial clustering of applications with noise (DBSCAN) algorithm [6] to cluster the data and detect anomalies sample points in smart buildings. We hypothesize that the clusters or groups that contain a large number of the time segments represent nominal operations, whereas outliers (the objects that are not reachable from any other object in the DBSCAN) and smaller clusters may represent anomalous situations. In previous work, Li et al [11] have used DBSCAN for anomaly detection.

DBSCAN considers an object o a core object if at least $MinObj$ objects are within its ϵ distance (including o). $MinObj$ and ϵ are input parameters determined by users. An object q is said to be directly reachable from a core object o if $dis(o, q) \leq \epsilon$. An object q is reachable from a core object o if there is a path $p_1, ..., p_n$ with $p_1 = o$ and $p_n = q$, where each p_{i+1} is directly reachable from p_i (all the objects on the path must be core objects, with the possible exception of q). If an object is reachable from any object of the cluster, it is part of the cluster as well. All objects not reachable from any other object are outliers.

The DBSCAN algorithm automatically determines the clusters and the anomalies in the dataset. The time complexity of DBSCAN is mostly governed by finding the neighbors of each sample point, which in the worst case has $O(n)$ complexity, where n is the number of objects. Therefore, the complexity of the DBSCAN algorithm is $O(n^2)$. This time complexity can be reduced to $O(nlogn)$ by building more sophisticated data structures [10].

2.2.2 Discovering undetected faults and their significant features:

Steps 3 is designed to assign meanings or labels to the derived clusters to help the experts to identify undetected faults. As a first step, we consider large groups derived from the clustering algorithm to be nominal (this corresponds to the assumption that the system operates normally most of the time). The smaller clusters that are sufficiently distant from the nominal clusters are labeled as outliers. As discussed earlier, smart buildings are complex, and they may involve multiple phases and operational modes, corresponding to different operation times or environmental conditions. Some of the smaller clusters that are not detected by the model-based diagnoser, may, in reality correspond to special modes of operation, and, therefore, are not of interest in discovering discrepant and faulty behavior. Therefore, an additional challenge we face in this work is separating the special modes of operation from truly undetected faults.

We have developed an approach to extract additional cues to identify special operating modes. To facilitate discovery of anomalies, we defined *significant features* in a previous work [3] as the set of features that best differentiate each anomalous group from the nominal groups. These features help our human experts better understand and characterize the anomalous situation as potential faults, or special modes of operation. In previous work [3], we developed an Euclidean distance based measure to identify and extract significant features. To be consistent with the DBSCAN algorithm we redefine the significant features as follows in this paper.

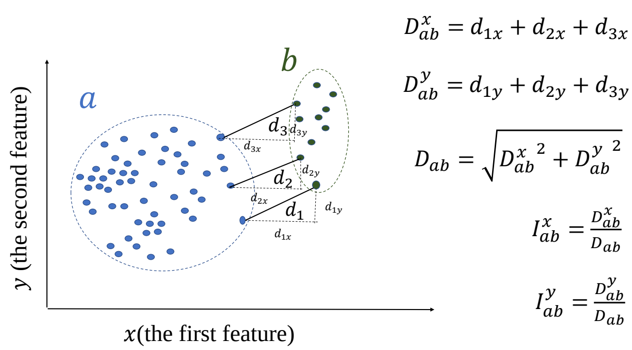

Figure 2: Importance of a feature.

Consider the k pair of objects with minimum distances in a way that each pair has an object in normal cluster a and an object in a small cluster b. Figure 2 shows an example where $k = 3$. In this case d_1, d_2 and d_3 are the 3 minimum distances between objects in a and b. We define the distance measure between normal operation cluster, a, and a small cluster, b, for signal variable, j, D_{ab}^{j}, as the sum of distances between the value of signal j in the k selected pairs. For example $D_{ab}^{x} = d_{1x} + d_{2x} + d_{3x}$, where d_{ix} is the distance between feature x in the ith pair in Figure 2. The total distance between cluster a and cluster b can be computed as

$$D_{ab} = \sqrt{\sum_{j=1}^{m} (D_{ab}^{j})^2}, \tag{5}$$

where m is the number of features in the data.

We define the importance of each feature v_j in distinguishing an outlier cluster, b, from normal operations, a, $I_{ab}(v_j)$, as the ratio of D_{ab}^{j} to D_{ab}, i.e.,

$$I_{ab}(v_j) = \frac{D_{ab}^{j}}{D_{ab}} \tag{6}$$

The importance of a set of features, $V_k = \{v_1, v_2, ..., v_k\}$ in distinguishing b from normal operation, a, is defined as:

$$I_{ab}(V_i) = \sqrt{\sum_{i=1}^{k} (I_{ab}(v_k))^2}. \tag{7}$$

Let $V = \{v_1, v_2, ..., v_m\}$ denote the set of features. We select a subset of features V_b to guarantee a minimum required importance, I_r, in distinguishing b from normal operation with minimum cardinality, i.e.,

$$\begin{aligned} \min \quad & V_b \subseteq V \\ \text{s.t.} \quad & I_{ab}(V_b) > I_r \end{aligned} \tag{8}$$

Once the significant features have been established, like model-based diagnosis, we can apply the Z-test to detect significant changes in the features. Consider V_s as the set of significant features for the system. We consider the last N_2 values to compute the mean value of a feature $v \in V_s$, $\mu_v(k)$, and the last N_1 samples ($N_1 \gg N_2$) to compute the feature variance, $\sigma_v^2(k)$ at sample time k. The confidence level for the Z-test, α, determines the bounds, z_-, and z_+, and, therefore, the sensitivity to the feature changes. The Z-test for feature v is:

$$m_v = \begin{cases} 0 & if \quad z_- < v(k) - \mu_v(k) < z_+ \\ 1 & otherwise. \end{cases} \tag{9}$$

where m_v represents the monitor derived from v, and M_v is the set of monitors derived from the significant features V_s.

2.3 Model-based + data-driven anomaly detection

A common approach to detect and isolate system faults is to derive a diagnosis reference model (see Figure 6) which is a bipartite graph that captures the causal relations between a fault (cause) and observed symptoms, which are deviant feature values. In traditional diagnosis systems, deviant feature values are captured in the form of monitors. In the model-based approach, the monitors are the outputs of the hypothesis tests that represent statistically significant changes in the residuals. In the data-driven approach, the monitors are the outputs of the hypothesis tests that represent statistically significant changes in the individual features or combinations of features. By considering the residual outputs and the significant features simultaneously, we combine model-based diagnosis and data-driven approach to achieve better fault detection and isolation. Therefore, we define the set of monitors for our hybrid diagnosis approach, M_h, as the union of the model-based monitors and data-driven monitors.

$$M_h = M_r \cup M_v. \tag{10}$$

In the diagnosis step, our goal is to use the set of monitors, $M_h = \{m_1, m_2, ...m_p\}$, to estimate the probability of each potential fault, $f \in F$, $P(f|M_h)$. The diagnoser uses the likelihood of each potential fault to rank the fault hypotheses and generate alarms. To estimate the probability of each potential fault a reasoning approach such as Bayesian Networks (BN) [2] can be applied. When the monitors are independent, we can make the Naive Bayes assumption and update the likelihood of a fault mode, f as follows.

$$P(f|M_h) = \alpha P(M_h|f) = \alpha P(m_1|f) \times ... \times P(m_p|f) \tag{11}$$

where α is a normalizing constant. As additional monitors fire, the number of likely fault hypotheses, defined by their likelihood of occurrence, become smaller, and with a lot of supporting evidence, may reduce to a most likely single fault. When more than one failure mode is likely, the reasoner ranks the active hypotheses in the order of their likelihood of occurrence.

In a previous work Mack et al [13] applied the Tree Augmented Naive Bayesian learning algorithm (TAN) to estimate the probability of each potential fault when the Naive Bayes assumption does not hold. The TAN structure provides a simple extension to the Naive Bayes model by capturing the dependencies among the monitors in the diagnosis reference model. Like the aircraft flight data [13], the monitors in the smart building data-set in our case study are not independent. Therefore, we use the Bayesian TAN classifier to update the likelihood of the fault modes in the next section.

3 Case study

The outdoor air unit (OAU) that we study in this paper was designed by Trane® and installed in Lentz public health center in Nashville to improve indoor air quality and reduce heating,

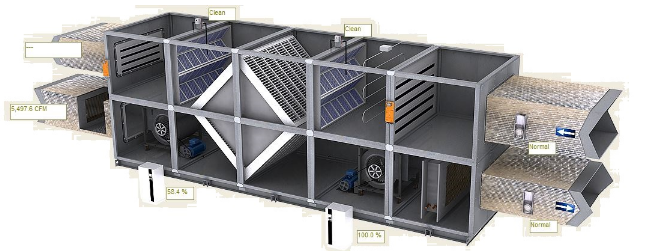

Figure 3: Output air unit (OAU).

ventilation and air conditioning (HVAC) costs. As it is shown in Figure 3, the OAU has two fans; the exhaust fan, and the outside fan. These fans control the building's input and output airflow in order to improve the indoor air quality, relative humidity and save energy by reducing the load on other HVAC equipments in the building.

3.1 The OAU model

It is challenging to model the OAU in sufficient detail to construct a model based diagnoser for the system. The relationship between a fan's static pressure and airflow has been shown to be nonlinear and a function of the fan's rotational speed[1]. The performance of the exhaust fan and the output fan are not independent. However, it is not trivial to mathematically model the effect of one on the other one. Finally, there are several system and environmental parameters such as wind speed, and the air filter's resistance that affect the model but are unknown.

Even though, deriving a complete model for the OAU system is not feasible, we can use physical laws to derive the relationships between the exhaust fan and the outside fan speed, static pressure and airflow. The law for incompressible flow applied to this configuration is given as:

$$\begin{aligned} Q_2 &= Q_1 (\frac{D_2}{D_1})^3 \frac{N_2}{N_1} \\ P_2 &= P_1 (\frac{D_2}{D_1})^2 (\frac{N_2}{N_1})^2 (\frac{\rho_2}{\rho_1}), \end{aligned} \tag{12}$$

where Q_i is the airflow in fan i, D_i is the diameter of fan i, N_i is the rotational speed of fan i, P_i is the static pressure in fan i, and ρ_i is the air/gas density in fan i.

3.2 The OAU system measurements

The OAU has 6 sensors that measure the statistic pressure in the output fan, the statistic pressure in the exhaust fan, the output fan rotational speed, the exhaust fan rotational speed, output fan airflow, and exhaust fan airflow. In addition to these continuous measurements, there are 14 sensors that record binary events such as the fire alarm status (on/off) or a fan filter status (clean/dirty). Therefore, the data we use for this case study contains 20 time-series waveforms. The sampling rate is 6 samples per hour and we have access to approximately 3 months of data. Overall, the data set contains 11193 samples. There are several data points

[1] https://www.trane.com/content/dam/Trane/Commercial/global/products-systems/education-training/continuing-education-gbci-aia-pdh/Fans-in-Air-Handling-Systems-/Trane_ENL_Fans_in_Air-Handling_Systems.pdf

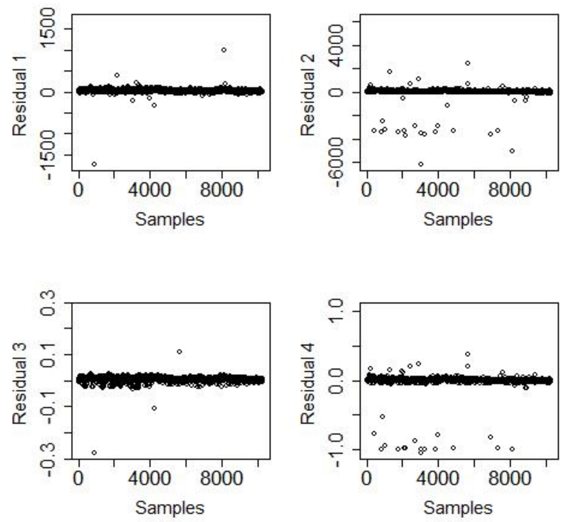

Figure 4: Value of residuals.

with missing values in the data-set. We remove these data points in the per-processing step. After removing the incomplete data points, each time series contains 10316 samples.

3.3 Model-based FDI

We can generate the following residuals to monitor the exhaust fan and the outdoor fan using the fan laws in (12).

$$\begin{aligned} r_1 &= Q_e - \sigma_e k_1 N_e \\ r_2 &= Q_o - \sigma_o k_2 N_o \\ r_3 &= P_e - \sigma_e k_3 N_e^2 \\ r_4 &= P_o - \sigma_o k_4 N_o^2, \end{aligned} \tag{13}$$

where Q_e is the exhaust fan airflow, N_e represents the exhaust fan rotational speed, Q_o is the outdoor fan airflow, N_o represents the outdoor fan rotational speed, and P_e and P_o represent the exhaust fan static pressure and the outdoor fan static pressure respectively. σ_e and σ_o are binary variables, where $\sigma_e = 1$ means the exhaust fan is on, $\sigma_e = 0$ means the exhaust fan is off, $\sigma_o = 1$ means the outdoor fan is on, and $\sigma_o = 0$ means the outdoor fan is off.

k_1, k_2, k_3, and k_4 are constants and can be computed using fan parameters, initial conditions, and air density or estimated using the historical data. In this paper, we use a maximum likelihood estimator to estimate these parameters. Table 1 represents the estimated parameters for the OAU system. Figure 4 shows the value of residuals for the estimated parameters. We consider 0.99 confidence level, α, for the Z-test. α defines the lower and upper bounds, z_- and z_+ in equation (4). Figure 5 shows the model based FDI results using the Z-test, where 1 represents abnormal behavior and 0 represents normal operation.

Figure 6 shows the diagnosis reference model using the set of residuals. We derived the residuals in equation (13) using the OAU system's steady state equations. Therefore, when the

Table 1: Residual parameters.

Parameter	Value
k_1	86.2834
k_2	91.84286
k_3	0.0001472726
k_4	0.0001595658

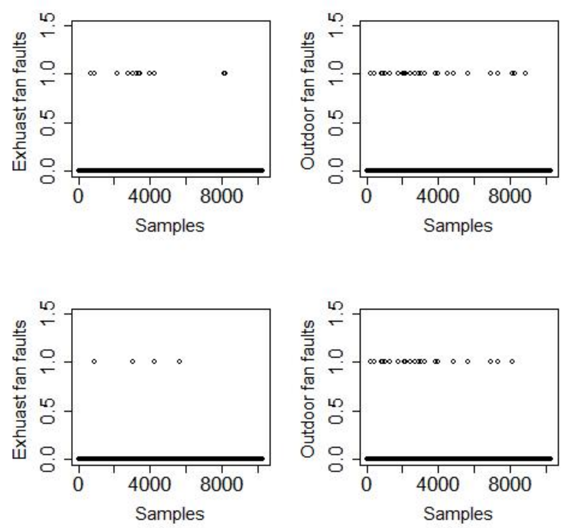

Figure 5: Hypothesis test outputs.

system is in a transition mode, it can trigger any of the derived residuals. Transition mode is not an abnormal behavior, and it is important to isolate this mode from potential abnormal behaviors. The first fault in the system, f_1, occurs when the exhaust fan or the outdoor fan is operating and the other one is off. The exhaust fan and the outdoor fan are supposed to always work together. However, it is possible that because of a software error, one of them starts operations earlier than the other fan. As we mentioned earlier, the performance of the exhaust fan and the output fan are not independent and when one of them stops, it affects the other one. We used normal operational data, where both of the fans work at the same time, to estimate residual parameters in Table 1. Therefore, when only one of them is working a number of residuals can be triggered. We show this in Figure 6 by connecting f_1 to all the residuals.

The fan filters can accumulate dirt as the system operates overtime. When a fan's filter becomes dirty, the resistance to airflow in the fan increases and the fan laws in equation (12) with the estimated parameters in Table 1 are not sufficiently representative of the observed behavior. Therefore, the exhaust fan filter status or the outdoor fan filter status can trigger the exhaust fan residuals (r_1 and r_3) and the outdoor fan residuals (r_2 and r_4) respectively. We can see that the diagnosis reference model in Figure 6 is not enough to isolate the transition mode and the faults. In the next subsection, we use a data-driven anomaly detection approach to

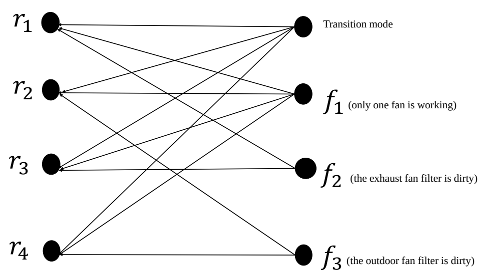

Figure 6: Diagnosis reference model using residuals.

extract significant features of each operation mode and fault to complete the diagnosis reference model.

3.4 Data-driven anomaly detection

3.4.1 Pre-processing:

When the measures are sensitive to the amplitude of the input signals, it is necessary to standardize the time-series waveforms before running a clustering algorithm to discover different modes of operation, including anomalous and faulty modes. Milligan and Cooper [15] performed an experimental study of seven standardization methods for clustering plus no standardization at all and concluded the approaches which standardize by division by the range of the variable give superior performance in recovering the cluster structure in the presence of noise. These approaches use the range of each feature to standardize it. A common method to standardize feature, F, is

$$F_s = \frac{F - \min(F)}{\max(F) - \min(F)}, \tag{14}$$

where $0 \leq F_s \leq 1$.

3.4.2 Choosing the DBSCAN algorithm parameters:

To the best of our knowledge, all of the well-known clustering algorithms require some input parameters. These parameters have a significant influence on the clustering result and typically, there is no unique approach to determine them. Unlike the k-means algorithm, DBSCAN does not require the number of clusters as an input. However, as it is mentioned earlier, it has two inputs: ϵ and $MinObj$. A common approach to determine ϵ is to compute the distance of k nearest neighbors of each object point for some k determined by the user and select ϵ where a sharp change is observed [7]. Figure 7 represents the K-dist plot for the OAU data-set for $k = 5$. The graph has a knee at $\epsilon = 0.02$. In this work, we are interested to identify small clusters as the possible anomalous groups. Therefore, we set $MinObj = 10$.

3.4.3 Clustering:

We apply the DBSCAN clustering algorithm (the R function, *dbscan*[2])) to cluster the OAU system data-set. Figure 8 shows that the algorithm generates six clusters. From the 10316

[2] see https://cran.r-project.org/web/packages/dbscan/dbscan.pdf

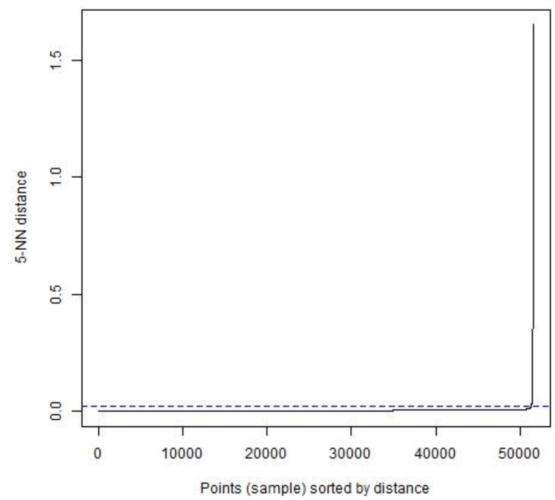

Figure 7: K-dist plot for the OAU system.

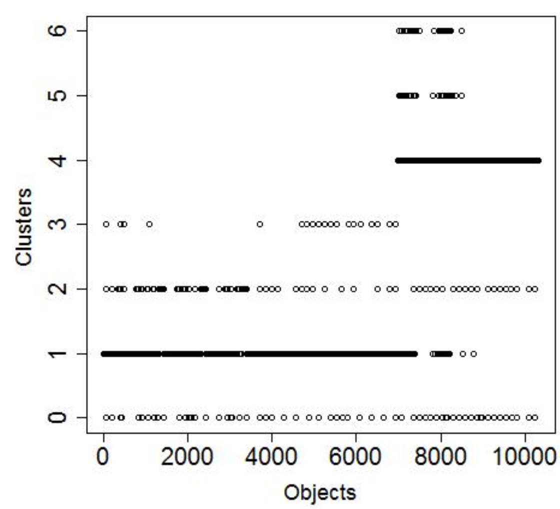

Figure 8: Clusters and outliers.

samples in the data-set, there are 64 outliers (cluster 0 in Figure 8), 6299 samples are grouped in cluster 1, 852 samples are grouped in cluster 2, 19 samples are grouped in cluster 3, 2968 samples are grouped in cluster 4, 55 samples are grouped in cluster 15, and cluster 6 has 59 samples. Cluster 1 is the largest cluster in the data-set and we assume that it represents the nominal behavior of the system. We study the other clusters in greater detail by comparing

Table 2: Summary description of the detected modes and anomalies

Cluster	Detected Mode or Anomaly	Significant Features	Description
1	Normal operation mode		
2	Mode: the system is off	• Exhaust fan status • Outdoor fan status	The OAU is off in this mode
3	Mode: transition	• Outdoor fan static pressure • Outdoor fan airflow • Outdoor fan speed command	Low pressure and airflow when the system starts
4	Fault: only one fan is working	• Exhaust fan speed command • Exhaust fan status	Exhaust fan off and outdoor fan on
5	Fault: the exhaust fan filter and the outdoor fan filter are dirty	• Exhaust fan filter status • Outdoor fan filter status	The exhaust fan filter and the outdoor fan filter have to be changed.
6	Fault: the outdoor fan filter is dirty	• Outdoor fan filter status	The outdoor fan filter has to be changed.
7	Fault: the exhaust fan filter is dirty	• Exhaust fan filter status	The exhaust fan filter has to be changed.

them against the nominal cluster to determine if they represent special modes or anomalies.

3.4.4 Extracting the significant features of each cluster:

To characterize the clusters, we extracted the significant features that differentiated each cluster from the nominal cluster. Using equation (8) we selected significant features by setting the threshold, $I_r = 0.9$. The significant features represented an ordered subset of features that contributed the largest amounts to the distance from the object in the outlier cluster to the object in the nominal with minimum distance, and the chosen subset accounted for 90% of

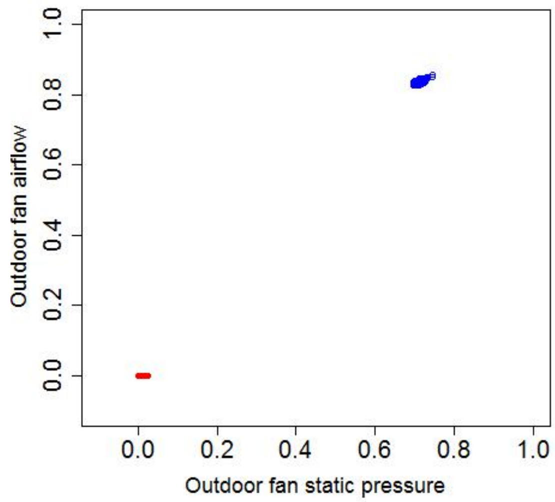

Figure 9: Outdoor fan airflow and outdoor fan static pressure in cluster 1 (blue) and cluster 3 (red).

the distance between the outlier and nominal cluster closest objects. Table 2 summarizes the significant features of six outlier clusters we identified by our clustering approach. We presented the significant features for each anomalous cluster to our experts to help us further characterize and classify the special modes and anomalies.

For example, consider cluster 3. In this cluster, the binary variables are the same with nominal cluster. However, as shown in Figure 9, the outdoor fan airflow and outdoor fan static pressure in cluster 3 are significantly different from the nominal cluster. The experts in the department of general services confirmed that this cluster represents the transition period, where the system is operating, but the airflow and the statistic pressure have not reached to their nominal values. We consider 0.99 confidence level, α, for the Z-test. α defines the lower and upper bounds, z_- and z_+ in equation (4). Figure 10 shows the Z-test results for the first two significant features of cluster 3, the outdoor fan static pressure and the outdoor fan airflow, where 1 represents abnormal behavior and 0 represents normal operation.

3.5 Fault diagnoser

In the first step, we use the significant features in Table 2 to complete the diagnosis reference model in Figure 6 as shown in Figure 11. Then we use the TAN classifier from the R package bnlearn[3] to train the reasoner. To validate the performance of our proposed approach, we implement a tenfold cross validation approach for two scenarios for our experimental study. In the first scenario, we use the diagnosis reference model derived from the residuals in Figure 6 and in the second scenario, we use the updated diagnosis reference model in Figure 11. Table 3 shows significant improvement in the average accuracy and the false positive rate when we use the combined model-based and data-driven approach.

[3] see https://cran.r-project.org/web/packages/bnlearn/bnlearn.pdf

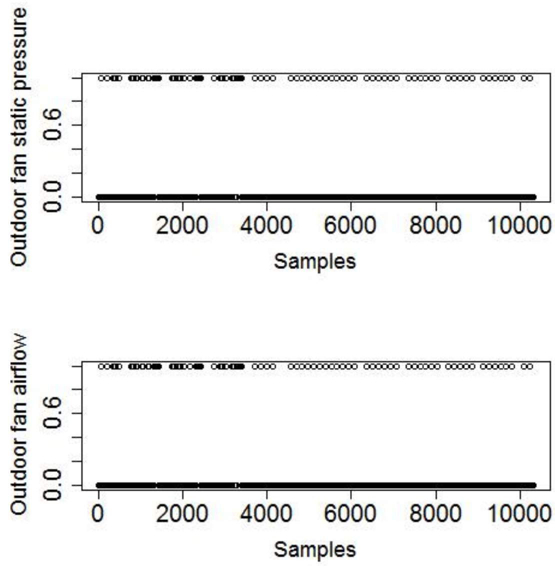

Figure 10: Hypothesis test outputs for outdoor fan static pressure and the outdoor fan airflow.

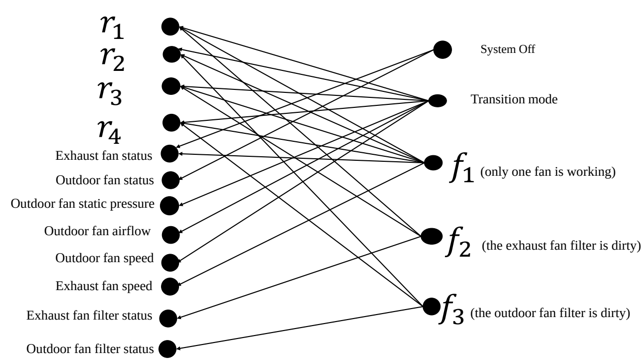

Figure 11: Diagnosis reference model using residuals and significant features.

Table 3: Accuracy and false positive rate for classifiers.

Diagnosis approach	Accuracy	False positive rate
Model-based approach	87.1%	12.7%
Hybrid approach	92.5%	2.5%

4 Conclusions

In this paper, we presented a methodology that combines model-based and data-driven methods to improve fault diagnosis performance in complex systems with incomplete models. Our approach applies an unsupervised learning method combined with human-expert support to update and complete the diagnosis reference model derived from the system model. We have described different steps of the method from the model-base FDI, data-driven anomaly detection, associating significant features with the outlier groups, generating diagnosis reference model, and designing classifiers. To demonstrate and validate the proposed method, we applied our hybrid approach to analyzing the data from a subsystem of Lentz Public Health Center in Nashville. The experimental results show the proposed hybrid approach significantly improves the diagnosis accuracy and reduces false positive rate.

Acknowledgements

The authors gratefully acknowledge the support provided by Freddie Adom, and Glenn Mohon in the Department of General Services at metropolitan government of Nashville in acquiring, analyzing, and interpreting the data. Hamed Khorasgani was partially supported by an internship from the Metro General Services Administration in Summer 2016.

References

[1] E. Alcorta-Garcia and P. Frank. Deterministic nonlinear observer-based approaches to fault diagnosis : A survey. *Control Engineering Practice 5.5*, pages 663–670, 1997.

[2] David Barber. *Bayesian reasoning and machine learning.* Cambridge University Press, 2012.

[3] Gautam Biswas, Hamed Khorasgani, Gerald Stanje, Abhishek Dubey, Somnath Deb, and Sudipto Ghoshal. An approach to mode and anomaly detection with spacecraft telemetry data. *International Journal of Prognostics and Health Management.*, 2016.

[4] Gautam Biswas, Gyula Simon, Nagabhushan Mahadevan, Sriram Narasimhan, John Ramirez, and Gabor Karsai. A robust method for hybrid diagnosis of complex systems. *In Proceedings of the 5th Symposium on Fault Detection, Supervision and Safety for Technical Processes, pp. 1125-1131.*, 2003.

[5] Anibal Bregon, Gautam Biswas, Belarmino Pulido, Carlos Alonso-Gonzalez, and Hamed Khorasgani. A common framework for compilation techniques applied applied to diagnosis of linear dynamic systems. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS*, 44(7), July 2014.

[6] Martin Ester, Hans-Peter Kriegel, Jorg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In Kdd, vol. 96, no. 34, pp. 226–231, 1996.

[7] Manisha Naik Gaonkar and Kedar Sawant. Autoepsdbscan: Dbscan with eps automatic for large dataset. *International Journal on Advanced Computer Theory and Engineering 2, no. 2, 11-16.*, 2013.

[8] Janos Gertler. *Fault detection and diagnosis in engineering systems.* CRC press, 1998.

[9] Daniel Jung, Kok Yew Ng, Erik Frisk, and Mattias Krysander. A combined diagnosis system design using model-based and data-driven methods. *In Control and Fault-Tolerant Systems (SysTol), 3rd Conference on, pp. 177-182. IEEE*, 2016.

[10] Amin Karami and Ronnie Johansson. Choosing dbscan parameters automatically using differential evolution. *International Journal of Computer Applications 91, no. 7.*, 2014.

[11] Lishuai Li, Santanu Das, R. John Hansman, Rafael Palacios, and Ashok N. Srivastava. Analysis of flight data using clustering techniques for detecting abnormal operations. *Journal of Aerospace Information Systems*, 12(9):587–598, 2015.

[12] Lishuai Li, Maxime Gariel, R. John Hansman, and Rafael Palacios. Anomaly detection in onboard-recorded flight data using cluster analysis. In Digital Avionics Systems Conference (DASC) IEEE/AIAA 30th, pp. 4A4-1, 2011.

[13] Daniel LC Mack, Gautam Biswas, Xenofon D Koutsoukos, and Dinkar Mylaraswamy. Learning bayesian network structures to augment aircraft diagnostic reference models. *IEEE Transactions on Automation Science and Engineering*, 2017.

[14] T. Soni Madhulatha. An overview on clustering methods. *arXiv preprint arXiv:1205.1117*, 2012.

[15] Glenn W. Milligan and Martha C. Cooper. A study of standardization of variables in cluster analysis. *Journal of classification*, 5(2):105–119, 1988.

[16] Pieter J Mosterman and Gautam Biswas. Diagnosis of continuous valued systems in transient operating regions. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 29(6):554–565, 1999.

[17] Sriram Narasimhan, Indranil Roychoudhury, Edward Balaban, and Abhinav Saxena. Combining model-based and feature-driven diagnosis approaches-a case study on electromechanical actuators. *21st International Workshop on Principles of Diagnosis*, 2010.

[18] Oliver Niggemann, Gautam Biswas, John S. Kinnebrew, Hamed Khorasgani, Sren Volgmann, and Andreas Bunte. Data-driven monitoring of cyber-physical systems leveraging on big data and the internet-of-things for diagnosis and control. *26th International Workshop on Principles of Diagnosis, Paris, France*, 2015.

[19] Gregory Provan. Generating reduced-order diagnosis models for hvac systems. In Itnl. Workshop on Principles of Diagnosis, Murnau, Germany, 2011.

[20] Nida Sheibat-Othman, Nassim Laouti, Jean?Pierre Valour, and Sami Othman. Support vector machines combined to observers for fault diagnosis in chemical reactors. *The Canadian Journal of Chemical Engineering 92, no. 4, 685-695*, 2014.

[21] James M. Sinopoli. *Smart buildings systems for architects, owners and builders*. Butterworth-Heinemann, 2009.

[22] Heidar A. Talebi, Farzaneh Abdollahi, Rajni V. Patel, and Khashayar Khorasani. Neural network-based state estimation of nonlinear systems: application to fault detection and isolation. *Vol. 395. Springer*, 2009.

[23] Venkat Venkatasubramanian, Raghunathan Rengaswamy, Surya N. Kavuri, and Kewen Yin. A review of process fault detection and diagnosis: Part iii: Process history based methods. *Computers and chemical engineering 27, no. 3*, pages 327–346, 2003.