Comprehensive Failure Diagnosis Model with Degradation Indicators of Multiple Sensors

Jun Tominaga[1], Shoya Kamiaka[1], and Kohei Kuroda[1]

[1] *Technology and Innovation Center, Daikin Industries, Ltd. 1-1, Nishi-Hitotsuya, Settsu, Osaka, 565-8585, Japan*
jun.tominaga@daikin.co.jp
shouya.kamiaka@daikin.co.jp
kouhei.kuroda@daikin.co.jp

ABSTRACT

Although monitoring system can detect abnormality of sensor reading in air-conditioning equipment, the root cause of the abnormality may not be sensor failure but other failures such as gas shortage. We propose new method that estimates the cause by the following steps. Firstly, regression model predicts the normal readings of multiple sensors (e.g., thermistor) for a given operational condition. Secondly, the gap between measured and predicted values is calculated for each parameter as a degradation indicator. Finally, our failure diagnosis model estimates the cause by considering degradation indicators of multiple sensors. Our evaluation verifies the effectiveness of our method.

1. INTRODUCTION

Reducing downtime of hearing, ventilating, air conditioning, and refrigeration (HVAC&R) units is of crucial importance. One way to realize reduced downtime is to accurately diagnose the failed system and remove the root cause as soon as possible. Skilled engineers can complete the diagnosis accurately. However, it is unrealistic to secure a sufficient number of skilled engineers. By introducing a system that automatically and accurately diagnoses failures, therefore, it becomes possible to complete the repair at the first visit and reduce workload on engineers, not demanding sophisticated skills for engineers. Furthermore, more accurate diagnostics and systems that can identify correct spare parts in need will make possible to automatically arrange for spare parts, enabling swifter repairment. This paper proposes a new method for Fault Detection and Diagnosis (FDD) purposes described above.

Generally, an error code is often issued from failed HVAC&R systems to indicate abnormality that has occurred to them. These error codes can be informative for FDD. For some error codes, however, it is difficult to identify the root cause. For example, in the case of an error code indicating refrigerant leaks from HVAC&R system, where a significant amount of gas has leaked from units, it is not possible to determine whether the refrigerant is really decreased or sensor that measures the refrigerant level is faulty. In addition to error code issued by equipment, therefore, it is necessary to jointly use FDD system to achieve more accurate and precise diagnosis.

FDD is often realized with rules (Katipamula et al., 1999), which have been widely used for decades. One of the advantages of this approach is its high interpretability of decisions made by algorithm. Being able to incorporate expert knowledge into algorithm is another advantage to note. Moreover, rules can be created with a small amount of data, or even without any data, unlike machine learning or deep learning methods described in the rest of this chapter. On the other hand, rules also have some potential drawbacks. For example, actual HVAC&R units have too diverse operation conditions and failure modes to create optimized rules for each state. Furthermore, actual sensor readings are sensitive to ambient temperature, installation environment, and control conditions. Thus, it is difficult to create rules to isolate faulty operations for all patterns. Because rules reflect expertise of engineers, moreover, they are not easy to maintain. Indeed, it is necessary to re-create and adjust rules every time new model of HVAC&R is launched. Katipamula et al. (2005) collectively summarized advantages and disadvantages of rules including those described above.

Recently, Machine Learning (ML) and Deep Learning (DL) approach becomes more prevalent in FDD activity (Chakraborty et al., 2019; Tun et al., 2021; Zhang et al., 2023). In line with remarkable progress on ML and DL techniques, this trend led to dramatic increase in the amount of data available from HVAC&R units with the help of IoT devices. Data-driven approaches that combine large amounts of data with ML are capable of detecting and classifying faults with high accuracy. On the other hand, one of the disadvantages of such approaches is low interpretability of decisions made by algorithms. Although ML outperforms in extracting important features for diagnosis from large amounts of data, these features are not necessarily explanatory enough to users, which may hinder these approaches from being widely acceptable to them. In addition, data of failed units, which is indispensable for developing ML algorithms, can be limited. Since failure patterns of HVAC&R units are diverse and only small fraction of them actually fail in the market, it may be difficult to secure enough data to train algorithms for all failure modes.

This paper proposes a hybrid method that combines rules and ML to overcome issues described above. The method consists of two steps: (1) Prediction of normal sensor readings for a given HVAC&R unit with ML approach, and (2) rule-based failure mode classification focusing on deviation of actual sensor readings from predicted ones.

In step (1), firstly a ML model is trained with actual operation data of various HVAC&R units in the market. Secondly, the model is used to predict normal values of sensors for a given installation environment and control conditions. Since the normal value prediction model in this method learns a large number of HVAC&R data in the market, it can predict normal values for various ambient temperature conditions, installation environments, and control conditions. Previous studies (Mirnaghi et al., 2020) suggested that data-driven approach can detect faults more accurately for complex physical systems including HVAC&R units than conventional methods using physical models.

In step (2), the difference between the predicted value in step (1) and actually measured value of sensor is derived. This gap is defined as "the degree of abnormality" of the sensor. Next, rule-based classification of failure modes is performed focusing on this degree of deviation. These rules are set based on domain knowledge of HVAC&R units. Specifically, different failure mode results in different combination of these degrees of deviation, which can be used inversely to classify failure mode. One advantage of our hybrid method is that ML model with the help of a large amount of normal data mitigates differences among HVAC&R models. Moreover, using a simple indicator, the difference between actual measurement and prediction to determine failure modes, enables high interpretability of the diagnosis results. The fact that normal data is easy to obtain, unlike failure data, also enhances the applicability of this method.

Our organization has already applied this approach that focuses on the discrepancy between actual and predicted normal value with ML model, in order to detect thermistor degradation and refrigerant leak (Kimura et al., 2022). This paper extends this method further to classify not only specific faults such as thermistor degradation or refrigerant leak, but also various failure modes based on the degree of discrepancy. This paper describes an integrated diagnostic system for commercial HVAC&R units that calculates the degree of deviation of multiple variables and accurately diagnoses failures by utilizing these multiple indicators.

This paper is organized as follows. Section 2 is devoted to describing our method in a sequential manner; data preparation, model training, and performance evaluation. Section 3 summarizes the result of fault mode classification and demonstrates the advantages of our model. Section 4 discusses the limitation of our model and possible extension for wider applicability. Section 5 is to summarize this work.

2. METHODS

The procedure to verify our approach consists of 5 steps.

1. Extracting data from data sources
2. Training normal value prediction models
3. Creating FDD model with rules
4. Diagnosing failures
5. Evaluating diagnosis result and performance

The overall flow is shown in Figure 1.

2.1. Extracting data from data sources

In this paper, we choose commercial air conditioner (AC) among HVAC&R units as a target for FDD. Generally, more than 20 sensors are installed with commercial ACs, and their readings are collected on a regular basis as operating data. The proposed method uses operation data of commercial ACs in the market, error code history issued by ACs, and their repair history by engineers. Operating data is in the form of time-series for various thermistors, pressure sensors, expansion valves, compressors, control modes, and so on. Error code history is used to roughly determine whether AC of interest is normal or faulty. Repair history describes actual work conducted by engineer. By utilizing this data, it is possible to identify the root cause of failure of faulty equipment, which serves as correct label in supervised training of ML model as described below. This label is also used to evaluate classification accuracy of our FFD algorithm.

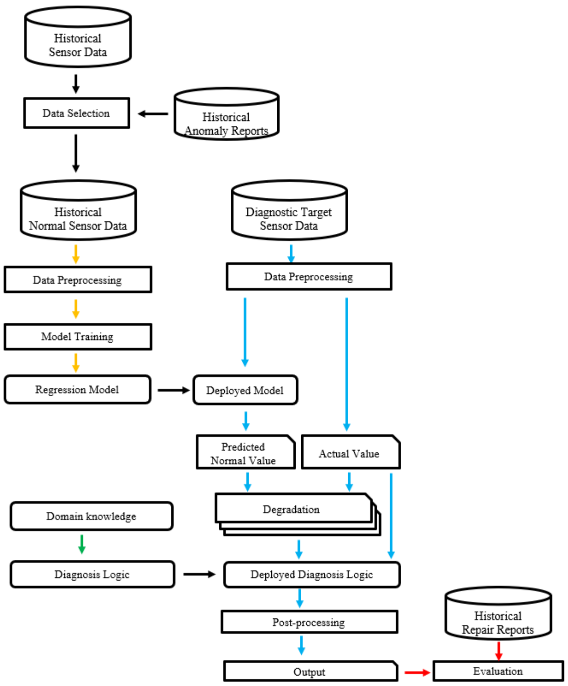

Figure 1. Process flow of our FDD system

2.2. Training normal value prediction models

Figure 2 shows how our normal value prediction model works. Blue line corresponds to actual sensor reading, while orange line describes predicted values by our model assuming the unit is not faulty. Comparison of these two values gives us clear insight on "abnormality level" of the unit, because they go more divergent as degradation of the unit proceeds (red arrow in Figure 2).

Only normal units which have not issued any error codes ever are used to train the normal prediction model. The error code history is used to classify normal and faulty units. In the training of models, the following data preprocessing is performed.

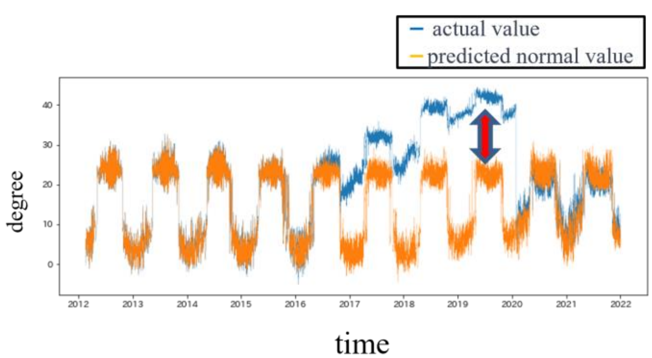

Figure 2: Schematic diagram of normal value prediction model

2.2.1. Selection of data items for training

Although various sensor data is collected in the market operation data, variables used as inputs to train the normal value prediction model are carefully selected to those considered important for system's refrigerant characteristics.

2.2.2. Derivation of explanatory variables

Some variables are more useful when combined with other variables for training. For example, it is known that Discharge Super Heat (DSH), which measures the temperature increase after compression, is an important variable to describe the state of refrigerant circuit. DSH is calculated discharge temperature measured from high pressure saturation temperature. Similarly, we convert some raw variables to combined ones for more efficient training.

2.2.3. Selection of equipment operating conditions

Since sensor values may not be stable just after AC is turned on or cooling (heating) mode is switched to heating (cooling) mode, we remove data obtained in such a transient state. Therefore, we use only steady-state operation data not only for training but also for diagnosis, assuming it contributes to reduce the chance of misjudgment in FDD.

2.2.4. Model training

Normal value prediction model is trained for each sensor. When training each model, we let a sensor to predicted be target variable, while the other sensors be input variables. For example, when we train a normal prediction model for thermistor A, sensors other than thermistor A are used as input variables to predict thermistor A. This corresponds to train the prediction model for thermistor A's normal readings for a given operation conditions (i.e., the other sensors), learning their mutual correlation behind. Besides, normal prediction model is trained separately for each operating mode (e.g., cooling or heating) as correlation among refrigerant to be learned varies depending on operation mode. In addition, training is performed for each AC model, as the number of sensors, their location, and correlation among sensors differ significantly over AC models. Several dozens of ACs in the market are used to train normal value prediction model for each AC model. The length of data for training is set to be one year to include all seasons. After pre-processing steps above, the value normal prediction model is trained with ML. We use LightGBM (Ke et al., 2017), a gradient boosting method, for creating models, where a sensor to predict as output variable and the other sensors as inputs.

2.3. Creating FDD model with rules

The deviation of actual sensor reading from those predicted by ML model is defined as the degree of abnormality for that sensor. The proposed method calculates deviation for all variables which are considered important for refrigerant characteristics and uses them for diagnosis. By making diagnosis on deviation, not sensor readings themselves, FDD algorithm can be common over different AC models. This means that it is no longer necessary to develop rules for each AC model, which improves the maintainability of FDD algorithms.

The FDD logic is expressed by rules based on the degree of deviation. These rules are created from domain knowledge on HVAC&R systems and adjusted to maximize the overall performance of diagnosis, whose indicator is described in subsection 2.5. Table 1 provides an idea on the FDD logic. For example, if an AC unit has a deviation for sensor1>a and sensor6<b, FDD logic diagnoses gas shortage as the root cause of failure. The FDD logic can be created for each failure mode. This table can propose more than one failure when multiple failures happened simultaneously, as long as an AC unit satisfies those conditions (e.g., broken suction thermistor and broken deicer thermistor).

Table 1: FDD logic based on deviation

Failure	Deviation for each sensor								
	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5	Sensor 6	Sensor 7	Sensor 8	Sensor 9
Gas shortage	>a					<b			
Broken suction thermistor		>c				<d			
Broken deicer thermistor			>e					>f	
Broken subcool heat exchanger gas pipe thermistor		>g							<h
Degraded compressor					>i		>j		

2.4. Diagnosing failures

With normal value prediction model and FDD table, we execute FDD for AC units of interest. Firstly, the same preprocessing as described above is executed to operation data to be diagnosed. The predicted value is derived for all sensors by making use of normal value prediction model trained above. Secondly, the difference between actual measured value and predicted value is calculated as the degree of deviation for all sensors. Thirdly, the deviation and measured values are input into the FDD logic to isolate the failure mode. In what follows, the post-processing to determine final output is outlined.

1. While the degree of deviation is calculated for each data point in the operating data, our approach does not directly use them for final output. This is because noise on derived deviation can be present, and it can lead to sporadic large deviation for normal sensor. This is partly due to insufficient training of normal units and should be addressed. Unfortunately, noise and deviation due to failure can be difficult to separate, because both of them appear as "deviation from normal". In order to prevent misdiagnosis due to noise, final output of FDD is made over a certain time range as a group. To be more specific, our FDD concludes failure mode only when large deviation is confirmed for a certain fraction of datapoints over a couple of days. This helps FDD correctly distinguish true failures from noises, because deviation by failures basically keeps growing unlike noises, which could suddenly be back to normal state. Therefore, FFD over a range of datapoints is of great use.
2. In the cause of multiple faults confirmed, furthermore, total duration of faulty conditions is displayed in order to prioritize failures to be fixed. For example, duration of faulty operations as shown in Table 2 makes

possible to list more than one faults, and the level of each abnormality can be expressed by their duration.

Table 2: Duration of faulty modes

No	Failure Mode	Duration
1	Failure Mode A	32 hours
2	Failure Mode B	3 hours
3	Failure Mode C	1 hour

2.5. Evaluating diagnosis result and performance

In this subsection, we evaluate the performance of our FFD approach with test dataset. We choose fault candidate with longest duration as output for this evaluation if multiple faults are detected. As the labels of failure are required to evaluate the performance of our model, repair history of engineers is used for that purpose. Label referred here corresponds to the spare part of AC that is replaced by engineer, such as suction thermistor, high pressure sensor, outdoor electronic expansion valve, compressor, gas shortage, and so on.

Table 3: Definition of confusion matrix

		Predicted	
		Normal	fault
Actual	Normal	*TN* (True Negative)	*FP* (False Positive)
	Fault	*FN* (False Negative)	*TP* (True Positive)

In this evaluation we focus on how accurately our FDD model can identify the failure modes for given failed units. After confirming that normal units in test dataset are not classified as failure, we then perform N-class classification for failed units in test dataset. As evaluation indices, we calculate precision, recall, and f1 score for each failure mode. Each indicator is defined as follows.

$$precision = \frac{TP}{TP + FP} \tag{1}$$

$$recall = \frac{TP}{TP + FN} \tag{2}$$

$$f1\ score = \frac{2 \times\ precision \times recall}{precision + recall} \tag{3}$$

3. Result

We investigate the performance of proposed FFD model to classify various failure modes by using discrepancy for all sensors. Specifically, we calculate precision, recall, and the f1 score. Each indicator is calculated for each failure mode. Table 4 shows the confusion matrix for one failure mode: gas shortage. In this case, gas shortage is treated as positive, and the other failures are defined as negative.

Table 4: Confusion matrix in diagnosing gas shortage

		Predicted	
		The others	Gas shortage
Actual	The others	78	0
	Gas shortage	4	16

Table 5 summarizes the performance for 5 failure modes, for which >5 failed samples are found in dataset. The f1 score is 0.8 or higher for all five failure modes. These results suggest high performance of the proposed method, though further study with more samples is required.

Table 5: Performance of FFD model

Failure	precision	recall	F1 score	Sample size
Gas shortage	1.00	0.80	0.89	20
Broken suction thermistor	1.00	0.94	0.97	16
Broken deicer thermistor	1.00	1.00	1.00	19
Broken subcool heat exchanger gas pipe thermistor	1.00	0.83	0.91	6
Degraded compressor	1.00	0.67	0.80	15

The method guarantees high interpretability with FDD results. To demonstrate that the pattern of deviations in failed data is consistent with the physics of the refrigerant cycle, we visualize both actual sensor readings and predicted normal values of a unit in gas shortage in Figure 3. Top panel refers to discharge pipe temperature, while bottom panel is expansion valve. In this case, measured values are greater than predicted normal values both in discharge pipe temperature and expansion valve, which is typical for refrigerant circuits in short of gas. As the behavior of the discrepancy is quite consistent with the physics of the refrigerant cycle, we conclude that our FDD model is highly interpretable in its reasoning.

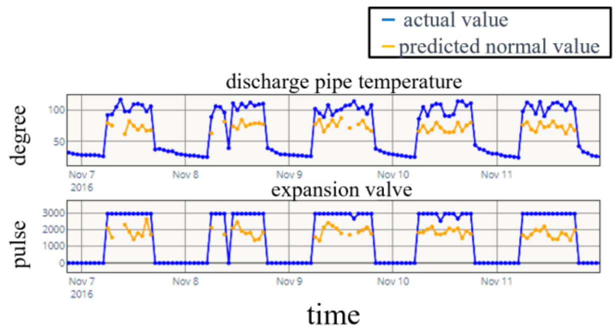

Figure 3: Degree of deviation of a unit in gas shortage

4. Discussion And Future Prospects

The method proposed in this paper is demonstrated to be capable of identifying the root cause of failure with high accuracy for five failure modes (gas shortage, broken suction thermistor, broken deicer thermistor, broken subcool heat exchanger gas pipe thermistor, and degraded compressor). Because the eligibility of our method is currently limited, however, further work to extend its applicability to more failure modes and HVAC&R models is needed. In what follows we refer to possible extension of this model.

4.1. Expansion of target models

In this study, the target of our approach is limited to a few AC models, not all models in the market. Since the proposed method uses deviations as input for FDD algorithm, this method is expected to be common over AC models, unlike rules, which are basically needed to be developed for individual models. Therefore, study to expand target models is highly promising.

4.2. Expansion of failure modes

The failure modes examined in this study do not include all failure modes, although they are mostly major failure modes in real-world operation. In this paper we do not study the rest of failure modes simply because of insufficient number of samples for such modes. Extending eligibility of failure modes with more samples in the market can make our model more comprehensive and reliable.

5. Summary

Accurate FDD when HVAC&R unit failed is fairly important for both customers in terms of reduced downtime of the units and engineers in terms of reduced man-hours for addressing failures. Although HVAC&R unit itself issues error codes in the event of failure, they may not be sufficient to identify the root cause of the failure. A common and useful FDD algorithm is rules, but they also bring some drawbacks such as maintainability.

Therefore, in this paper, we propose a new FDD algorithm that focuses on the gap between actual sensor readings and those predicted by ML model for a given operation condition. We verify its performance for commercial ACs. Specifically, our model can correctly separate normal and faulty units, and classify 5 failure modes (gas shortage, broken suction thermistor, broken deicer thermistor, broken subcool heat exchanger gas pipe thermistor, and degraded compressor) with high accuracy. As a whole, the diagnostic accuracy of the proposed method is high enough to correctly isolate failed parts. Moreover, the proposed FDD logic has high interpretability for the diagnosis results. In future work, we expand target models and failure modes with more samples and improve overall diagnosis accuracy to make our model more comprehensive and reliable.

References

Chakraborty, D., & Elzarka, H. (2019). Early detection of faults in HVAC systems using an XGBoost model with a dynamic threshold. *Energy and Buildings*, 185, 326-344. https://doi.org/10.1016/j.enbuild.2018.12.032

Katipamula, S., Pratt, R. G., Chassin, D. P., Taylor, Z. T., Gowri, K., & Brambley, M. R. (1999). Automated fault detection and diagnostics for outdoor-air ventilation systems and economizers: Methodology and results from field testing. *ASHRAE Transactions*, 105(1).

Katipamula, S., & Brambley, M. R. (2005). Methods for fault detection, diagnostics, and prognostics for building systems—A review, Part I. *HVAC&R Research*, 11(1), 3.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *In Advances in Neural Information Processing Systems* (pp. 3149-3157).

Kimura, S., Moriwaki, M., Yoshimi, M., Yamada, S., Hikawa, T., & Kasahara, S. (2022). *Proc. of 2022 CLIMA REHVA 14th HVAC World Congress*. Rotterdam, Netherlands: ID1603.

Mirnaghi, M. S., & Haghighat, F. (2020). Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings*, 229, 110492. https://doi.org/10.1016/j.enbuild.2020.110492

Tun, W., Wong, J. K., & Ling, S. H. (2021). Hybrid Random Forest and Support Vector Machine Modeling for HVAC Fault Detection and Diagnosis. *Sensors*, 21(24), 8163. https://doi.org/10.3390/s21248163

Zhang, F., Saeed, N., & Sadeghian, P. (2023). Deep learning in fault detection and diagnosis of building

HVAC systems: A systematic review with meta analysis. *Energy and AI*, 12, 100235. https://doi.org/10.1016/j.egyai.2023.100235