Real-Time HVAC Sensor Monitoring and Automatic Fault Detection System

Ying Guo, Josh Wall, Jiaming Li and Sam West

Abstract Real-time monitoring of heating, ventilation, and air conditioning (HVAC) systems is crucial to maintaining optimal performance such as providing thermal comfort and acceptable indoor air quality, guaranteeing energy saving, and assuring system reliability. In a realistic situation, HVAC systems can degrade in performance or even fail due to a variety of operational problems, such as stuck open or closed air dampers and water valves, supply or exhaust air fan faults, hot or chilled water pump faults, and inefficiencies in the way HVAC systems or pieces of equipemnt are controlled. This paper presents automatic fault detection techniques, as well as a key sensor sets selection approach that can help to maintain the performance of HVAC systems, and optimise fault detection results. One key step to make sure the approach succeeds is the sensor feature selection process. This paper implements the ensemble rapid centroid estimation (ERCE) as the data-driven sensor and feature selection algorithm, which is the core method to assure the automatic fault detection can function correctly. Instead of choosing sensors manually, ERCE method can automatically select representative features that are unique and relevant to the faults in a HVAC system. The methodology presented is implemented in real-world commercial buildings with experimental results showing that different types of faults are detected successfully.

1 Introduction

Heating, ventilation, and air conditioning (HVAC) systems are designed to provide thermal comfort and acceptable indoor air quality in a range of commercial buildings [1]. HVAC systems consume a large amount of energy throughout the world. For

Y. Guo (✉) · J. Li
Data61, CSIRO (Commonwealth Scientific and Industrial Research Organisation),
Marsfield, NSW 2122, Australia
e-mail: ying.guo@csiro.au

J. Wall · S. West
Energy, CSIRO (Commonwealth Scientific and Industrial Research Organisation),
Mayfield West, NSW 2304, Australia

example, in Australia, it is estimated that the installed base of non-residential HVAC systems consumes 9 % of electricity produced, representing more than 3.6 % of Australia's greenhouse gas emissions; and they create more than 55 % of electrical peak demand in commercial business district (CBD) buildings [2]. The Australia government recognizes that large efficiency gains can be achieved through ongoing maintenance and more optimal operation of HVAC systems in existing building stock, and seeks to establish national system standards of documentation for design, installation, operation and maintenance of HVAC equipment/systems.

As well as the mechanical and electrical components and pieces of equipement, a HVAC system comprises some form of control logic to regulate the operation of the components and system as a whole. Usually a sensing device is used to compare the actual state (e.g. temperature or humidity) with a target state. Then the control logic draws a conclusion of what action has to be taken (e.g. provide more heating or cooling). Modern HVAC systems not only have fundamental sensors and actuating devices in addition to some basic control logic to perform their required function, but often include a more advanced building management and control system (BMCS) that provides multiple levels of control, data monitoring and analytics, user interfaces and even interfaces to other building energy systems.

A variety of sensing devices (such as temperature, humidity, velocity, or pressure) are installed in the HVAC systems. Sensors measure the actual value of a controlled variable such as temperature, humidity or flow and provides information to the BMCS. In a realistic situation, the building HVAC system can fail to satisfy performance expectations envisioned because of problems caused by improper installation of sensors, inadequate maintenance, and equipment or sensor failures. These problems, or "faults," include mechanical failures such as stuck, broken, or leaking valves, dampers, or actuators; control problems related to failed or drifting sensors, poor feedback loop tuning or incorrect sequencing logic; fouled heat exchangers; design errors; or inappropriate operator intervention. Such faults often go unnoticed for extended periods of time until the deterioration in performance becomes great enough to trigger comfort complaints, equipment failure or excessive power consumption.

Automated fault detection and sensor monitoring techniques for HVAC systems can identify these types of faults, with the potential energy-saving of avoiding these faults is estimated at 10–40% of HVAC system energy consumption, depending on the age and condition of the equipment, maintenance practices, climate, and building use [3–6]. By sensing and identifying minor problems before they become major problems, the useful service life of equipment can be extended. Also, repairs can be scheduled when convenient, avoiding downtime and overtime work. Depending on the building use, better control of the temperature, humidity, and ventilation rate of the occupied spaces can improve employee productivity, occupant comfort, and/or product quality control.

Most of the current commercially available solutions in HVAC sensor monitoring and fault detection systems use rule-based methods, where most solutions integrate and interpret incoming sensor data in accordance with the pre-determined set of rules, produce a risk profile, and initiate a response to a breach of these rules [7–10].

Another class of solutions use model-based methods, which use analytical mathematical models to compare and identify faults based on the sensor data sets [11–13].

Since every building is unique, it is not a simple task to set these rules or to generate these analytical mathematical models. In addition, the task of setting these thresholds used by such solutions to raise alarms is quite involved, and prone to producing false alarms. Hence, we developed a statistical machine learning algorithm based approach for the automated monitoring and fault detection in HVAC systems [14–17]. Our approach uses probabilistic models that are constructed on the probabilistic links between variables. Meanwhile, the probabilities are learnt from the stored sensor monitoring data sets. It is an ideal representation for combining prior knowledge and data, and can have much better flexibility and adaptability when applied to HVAC system.

For a complex HVAC system, the number of sensors and actuators can number in their thousands, and the selection of key sensors and actuators containing the main feature of the system and reflect to important system faults, is crucial for the success of our approach. This paper presents automated monitoring and fault detection techniques, and a key sensor sets selection approach to optimise the fault detection results. This methodology has been implemented and tested in real-world commercial buildings and experimental results show that different types of faults are detected successfully.

2 Overview on HVAC Systems, Sensor Monitoring and Fault Detection

A HVAC system normally includes central plant consisting of a hydronic heater, a hydronic chiller, a pump system, a valve system, a heat exchange system (which includes dedicated heated and chilled water coils), and an air distribution system for supplying occupants with conditioned air. It also includes a sensing system that includes a number of sensors located throughout the system, such as temperature, humidity, air velocity, volumetric flow, pressure, gas concentration, position, and occupancy detection sensors. The BMCS includes a computing system which interfaces with various sensory signals in the HVAC system. Using feedback from various components and sensors of the HVAC system, the environmental conditions for the inhabitancy or functional purpose of the building can be regulated. Figure 1 shows a simple schematic of a HVAC system. It consists of three main parts: air handling unit (AHU), the chiller (cooling) and boiler (heating) systems, and the control system. When a HVAC system is operational, a supply air fan as part of the AHU draws air from either outside, return air from the indoor area, or a mix of both, and past the cooling/heating coil heat exchange to achieve the desired temperature and humidity before being supplied to the indoor area. A trade off among exhaust, fresh and recirculation air is decided by the BMCS, based on the real-time sensor signals.
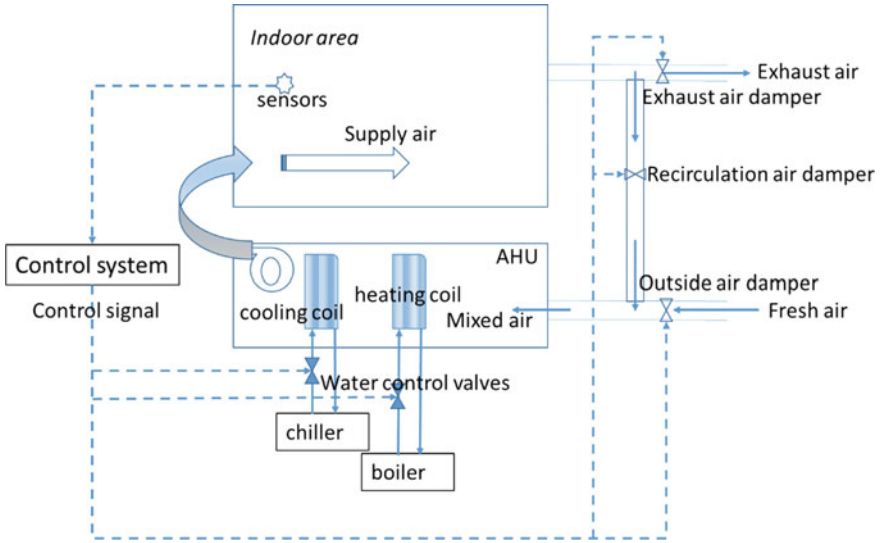
**Fig. 1** General schematic diagram of a typical HVAC system

In a typical multiple storey commercial building, there can be tens to hundreds of zones. A large high rise commercial office building needs to be divided into multiple zones, in order to satisfy and maintain desired temperature and air quality conditions. Figure 2 shows a screenshot of a single zone from a BMCS interface. Some sensors such as the damper positions, valve positions are marked in the figure, while some other sensors' readings are listed in the top right corner, such as the supply air temperature set point, return fan speed set point, etc.

Thousands of sensors read real-time status of the equipment in a large HVAC system. The abundance of sensor data makes it difficult and expensive for human operators to continuously monitor the system and identify faults or operational inefficienies quickly.

One solution is to develop an intelligent automated sensor monitoring and fault detection system which can continuously monitor sensor data from various system components and identify unusual or inefficient behaviours.

Our approach is to use statistic machine learning algorithms based on key sensor selection and monitoring technique. Firstly, historical sensor data is logged during normal operation of the HVAC system. Secondly, suitable sensor combinations and their features are chosen to train HVAC system status models. These self-learnt models can build up the time-varying relationships between monitored sensors and/or sensor features of normal operation in a HVAC system.

Finally, ongoing real-time sensor data is read in, and the likelihood of this data matching with learnt historical behaviour indicates whether the HVAC system is running as normal or not. Figure 3 shows the overview of this real-time monitoring and fault detection approach.
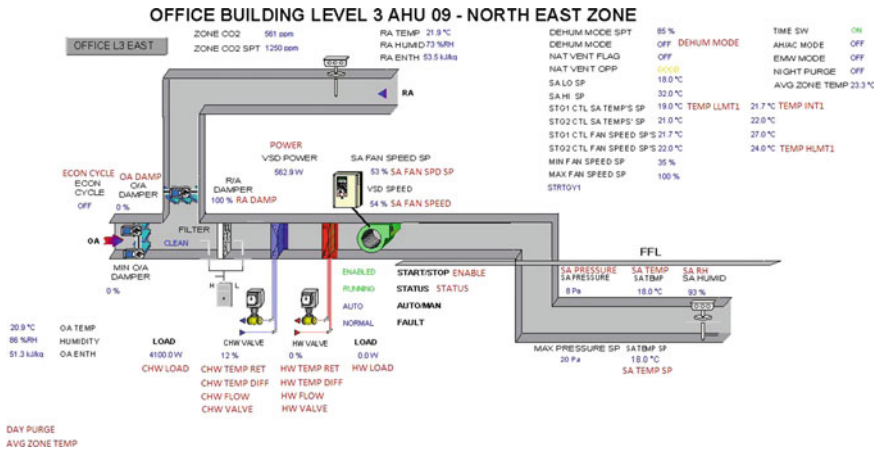
Fig. 2 The view of an Air Handling Unit (AHU) of a HVAC system from the Building Management & Control System interface.
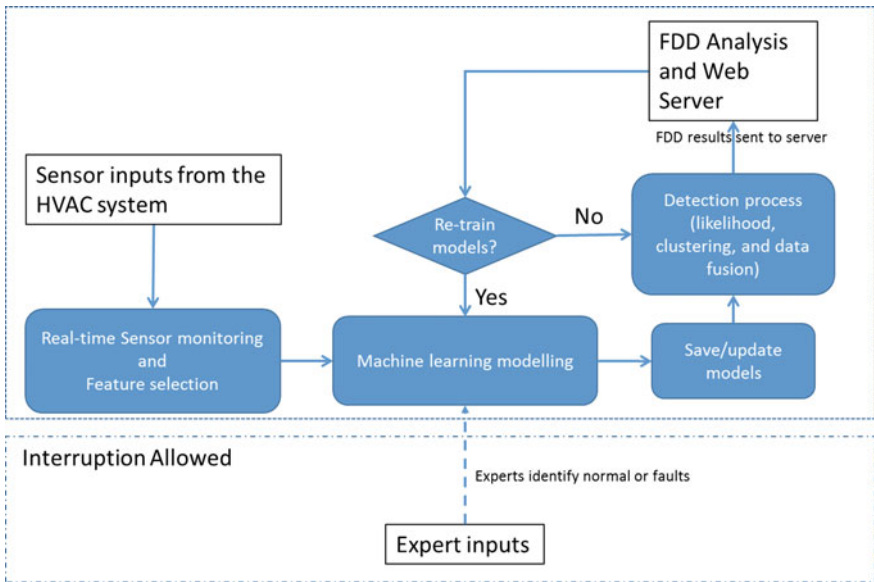
Fig. 3 Overview of the real-time sensor monitoring and fault detection approach

3 Sensor Monitoring and Fault Detection Approach in Detail

To overcome or ameliorate some of the limitations of existing fault detection methods, our approach is a combination of the state-of-the-art machine learning algorithms, including dynamic Bayesian networks (DBNs), Hidden Markov Models (HMMs), as well as swarm intelligence and consensus clustering methods. This section overviews our approach in details. Subsection A will present the key sensor selection for efficient real-time monitoring, the second subsection will show the architecture of the machine learning based methods, and the third subsection will list the main faults that are detected for HVAC systems.

Sensor and feature selection: As explained in Sect. 2, thousands of sensors are sending data to the database, and some of them are very crucial for proper system monitoring, modelling, and correct fault detection results. Proper sensor and/or sensor feature selection is essential for the whole model-based approach.

The aim of sensor selection is minimising redundancies between sensors so that the important system features are not undermined. HVAC system's performance may change dynamically depending on many conditions such as weather condition, seasonal condition, and occupancy of the building. Hence the sensors and their features need to be constantly monitored. Moreover, some sensors may contain little dynamic information, and can adversely affect the final model to an extent that some faults are missed. One way to decide the sensor combination is to depend on the HVAC technician's experience, but this is not an efficient way when it is applied to different structured buildings.

In our approach, we applied the rapid centroid estimation (RCE) as the key data-driven sensor and feature selection algorithm, which specifically performs well under varying seasonal conditions [18]. The feature extraction process from the sensor data involves statistical analysis [19–21] and dimensionality reduction [19, 21]. This is a crucial step, as inappropriate features could reduce the capability of the fault detection result.

We implemented an approach for sensor/feature selection using an ensemble clustering algorithm, which allows the natural recovery of clusters without having a priori knowledge regarding the optimum number of clusters. The method is Ensemble Rapid Centroid Estimation (ERCE) [22] based on the RCE algorithm [23]. ERCE exploits the fact that the quality of a clustering ensemble depends on the degree of diversity of the provided clusters. It shows better performance than conventional clustering algorithms such as complete linkage, ensemble k-means and ensemble fuzzy c-means.

The ERCE is a sequential process including clustering, fuzzification, and ensemble aggregation. In the clustering stage, various unique voronoi tessellations of the data is discovered; in the fuzzification stage, the voronoi tessellations are converted into fuzzy partitions; and in the ensemble aggregation stage, the final partition is recovered using the weighted fuzzy co-association-tree by hybrid

method [22, 23]. An overview of ERCE is presented as following, where a detailed description is available in [22, 23].

Given a data matrix Y,

$$Y = [y_1 \ldots y_j \ldots y_{n_j}],$$

where j denotes the observation index, n_j denotes the number of data (volume), and a particle position matrix X

$$X = [x_1 \ldots x_i \ldots x_{n_i}],$$

where i denotes the particle index, n_i denotes the number of particles, high dimensional voronoi tessellations are performed on the data such that each observation in Y is mapped to the nearest particle. In other words, each particle x_i governs a voronoi cell of the set C_i:

$$C_X = [C_1 \ldots C_i \ldots C_{n_i}], \; \emptyset \subseteq C_{1, \ldots, n_i}$$

which may contain empty sets. The clustered set,

$$C_X = C_{r, \ldots, n_c} \cap C_{i, \ldots, n_i}, \; \emptyset \subseteq C_X,$$

is the sets in $C_{i, \ldots, n_i}$ which partitions Y to n_c non-empty clusters.

The ERCE contains n_m swarms working in parallel such that

$$C_{\mathrm{ERCE}} = \left\{ C_{X_1} \ldots C_{X_{n_m}} \right\}.$$

Using the concept of charged particles [22], the possibility of creating duplicate partitions is minimised. Ideally each C_{X_m} would then return a unique partition of the data such that

$$C_{X_1} \neq C_{X_2} \neq \ldots \neq C_{X_m},$$

where each partition C_{X_m} denotes an optimal partition returned by the mth swarm.

After the clustering process, the label matrix is fuzzified based on the distance between particles and data $D = D(X, Y)$. The fuzzy membership value for the jth observation with respect to the ith cluster, u_{ij}, can be calculated as followings:

$$u_{ij} = \frac{e^{-d_{ij}/(2\lambda_i)}}{\sum_{i=1}^{n_i} e^{-d_{ij}/(2\lambda_i)}} \tag{1}$$

where d_{ij} is the distance between the ith particle to the jth observation, and λ_i denotes the ith bandwidth of the cluster centre. Here $\lambda_{1, \ldots, n}$ can be optimised using a compromise between the partition's fuzzified dissimilarity

$$D_{ij} = u_{ij}d_{ij},$$

and Shannon entropy

$$H\left(u_{ij}\right) = -u_{ij}\log u_{ij}$$

for each i and j. In other words, for each cluster, C_i, the optimum λ_i can be found by solving a convex optimisation problem:

$$\min_{s.t.\forall i, \lambda_i > 0} \|H - D\|^2, \tag{2}$$

which optimised λ_i for all cells, $i = 1, \ldots, n_i$, that the Gaussian probability distribution of the data governed in each corresponding voronoi cell is best described. In this approach, we use non-linear least square to optimise Eq. (2).

The examples in Sect. 4 will show that the feature selection is a powerful tool which can not only determine the feature cluster number, but also rank each feature within each cluster. The feature with the highest ranking in each cluster is then chosen as the key feature for fault detection.

Intelligent fault detection methods: The main processes of fault detection algorithm work as two steps, **training process step** (or called learning), and **fault detection step** (or called testing).

The historical sensor data which measures the **normal operation status** of the system was collected for **training process**. The intelligent sensor and feature selection process (as in Sect. 3A) is firstly implemented to prepare the training datasets. A statistical machine learning approach then learns the relativity between sensor measurements and system performance through these data. This approach uses probabilistic models that consist of variables and probabilistic links, which can denote the physical relationship between the sensor readings. Because of the complexity of the HVAC system, multiple models are learnt during the training process. The training process can be done during nights or weekends, and it normally takes about ten minutes over one week's datasets for one building. Because the system performance might change slowly because of season, weather, or other conditions, the training process can be repeated at a regular basis. For instance, the current training process can be done every week on tested buildings.

The **fault detection process** is to detect whether there is a fault in the system or not (binary classification problem). After the training process, a learnt normality model of the HVAC system can be used to detect system faults automatically. The collected sensor data is periodically detected with the models built based on the sensor features, and the similarity between current measurement and the historical feature is calculated. If the algorithm finds that the difference is obvious, a fault alarm is raised.

As the training process is a binary classification problem on time-series sensor features, we implemented a methodology which is based on a combination of the HMMs and the Support Vector Machines (SVM) algorithms [24]. HMMs are used

to denote the physical relationship between the sensor data in a dynamic system. Meanwhile, SVM can handle the nonlinear behavior very well, and requires small amount of training samples [16, 25, 26].

Furthermore, while the faults are detected as deviations in the normal operation, they can be the input data for training one or more of the fault detection models to learn patterns of faulty operation of the HVAC system wherein normal operation is detected as deviations in the faulty operation. As a dynamic process, the normality model and the fault models are adapted while the HVAC BMS dataset increases, hence the performance of real-time monitoring and fault detection improves.

The selected features of the sensor data are tested on different pre-trained HVAC normality models. Then the likelihood matrix is calculated as the indicator of the similarity between current sensor readings and normal system sensor readings. A couple of methodologies such as clustering and data fusion algorithms are also implemented as the final stage in the analysis of the likelihood matrix, and make the final decision as a sequence of binary value (e.g. Yes for normal, and No for abnormal). More details of the training methodologies are in [6, 15, 27].

Main faults for HVAC systems: It is reported that a few top listed faults are the major faults for HVAC systems [28]. Properly detect these faults can avoid up to 30% of energy waste in HVAC systems [29]. A short description of each fault is listed as following, and our sensor monitoring and fault detection approach will mainly focus on them.

- Hot water valve leaking or stuck: If the hot water valve stuck, hot air cannot be provided properly. If an internal valve leaks, it can be very difficult to troubleshoot and is often confused with a compressor that is not pumping to capacity. Both a leaking valve and a failing compressor have the same symptoms—both the heating and cooling capacity of the system are diminished. This is because the compressor continues to pump the air around and around inside the leaking valve.
- Supply-air fan belt slipping: The supply-air fan belt slipping can lead to the supply-air fan not running at setting speed.
- Outside air damper leaking or stuck: Damper not in the proper position. Outside air cannot circulate properly.
- Return air damper leaking or stuck: Return air cannot circulate properly.
- Individual zone temperature sensor fault: The temperature sensor reading in a zone is wrong.

4 Experimental Results

The intelligent real-time sensor monitoring and fault detection system has been tested on several buildings. In this section, we will show the results in one of these buildings, a large commercial building in Newcastle, Australia. The sensors are sending real-time data to the monitoring and fault detection system with a 1-min

interval which is typical of the resolution of trended sensor data from a BMS database [13]. Figure 4 shows an example of 15 sensors for an air handling unit (AHU).

Fault detection experimental results: During August and September 2013, 35 faults in four different fault types were generated for one AHU. Most faults last over six hours.

Table 1 listed the summary of the experimental results. Two types of faults, hot water valve stuck and individual zone air temperature failure are 100 % detected. But the success rate for the slipping supply-air fan belt fault is low. The lowest, though, is the return air/outdoor air damper point stuck fault, where only one third faults are detected. One of the main reasons is the measurement of the effects of each fault. For some faults, the return air qualities are affected directly, and corresponding sensor measurements are collected, such as return air temperature. For the return air damper, when it stuck at 70 %, the main facts that can change accordingly should be outdoor air fraction or indoor air CO_2. These facts were not measured or saved in the datasets for modelling.

Sensor monitoring and feature selection results: We have also analyzed the sensitivity of the sensor features for the fault detection in HVAC systems. The parameters that we judge include the data sample rate, training window size,

Fig. 4 An example of the sensor data for an air handling unit in the Newcastle building

Table 1 Summary of fault detection experimental results

Generated faults in Aug/Sep 2013	Detected	Not detected	Detection rate (%)	Comments
Hot water valve stuck	11	0	100	
Return air/outdoor air damper stuck	3	6	33	For the return air damper, when it stuck at 70 %, the main facts that can change accordingly should be outdoor air fraction or indoor air CO_2. These data were not measured or saved in the datasets hence not selected for modelling
Slipping supply-air fan belt fault	7	4	63.6	
Individual zone air temperature failure	4	0	100	

missing data tolerance, and minimal sensor subset performance. We chose the high frequency datasets for comparison experiments. The original dataset's sample rate is at one reading per 5 s, which is higher than most real HVAC systems. From the original trials, we observed that the training/testing window size is relevant to the sample rate. Hence we analyze the effects of the combination of these two parameters. We changed the sample rate to the following: 5 s, 20 s, 30 s, 40 s, 50 s, 60 s, 80 s, and 90 s. The training/testing window size changes between the following values: 10, 20, 50, 80, and 100. The combination of sensors is the same as default (six sensors).

Table 2 summarize the performance of the combination of these two parameters. Tick for successful detection, cross for unsuccessful detection. As we can see, when the sample rate increases, the window size need to be shorter to get the successful results. For the 90 s sample rate, the window size should be 20, which means that the training and testing data covers a 30–min period.

Table 2 Sensitivity of the sample rate versus window size of the sensor monitoring rate. Ticks for successful fault detection

| Sample rate → Window size ↓	5 s	20 s	30 s	40 s	50 s	60 s	80 s	90 s
10	×	×	×	×	×	×	×	×
20	×	×	×	√	√	√	√	√
50	×	×	√	√	√	√	√	×
80	×	√	√	×	×	×	×	×
100	√	√	×	×	×	×	×	×

Case study for sensor selection and its effect on fault detection: In this subsection, we will present one example to show the effect of sensor selection on fault detection results. The detail of this fault is as following:

- Date: 20/11/2009
- Location: All AHUs for Newcastle building
- Description: This is an actual fault that occurred when the chilled water system was under heavy load. Chiller 2 failed due to an overheating safety cut-out, and Chiller 1 failed to ramp up due to some misconfigured set-points, causing the chilled water temperature to rise to around 23 °, and in turn causing the chilled water valves on many zones (AHU9 included) to open up to near 100 %.

This type of faults is relatively common for HVAC systems, and can lead to a large waste of energy. The sensors that are selected to build the HMMs by the intelligent algorithms described in Sect. 3. A are: hot water valve sensor, chilled water valve sensor, supply air relative humidity sensor, supply air temperature sensor, return air relative humidity sensor, return air temperature sensor.

HMMs are trained on normal fault-free historical data. Then the real-time sensor monitoring data is fed in to calculate the likelihood. The lower the likelihood value, the higher possibility a fault exists.

The fault detection results are shown in Figs. 5 and 6. Figure 5 is the likelihood curves for the data on 20th Nov 2009, and the corresponding fault detection results after a classification and clustering process is shown in Fig. 6. It is clear that the fault, occurring after 12:00 pm, is convincingly detected, as shown by the likelihood curves.
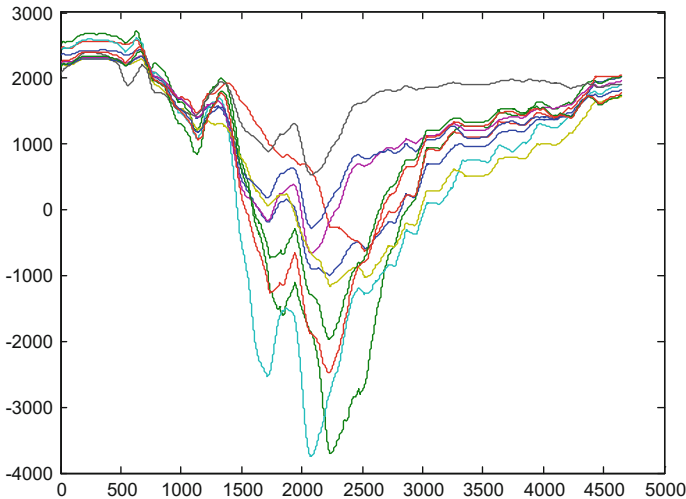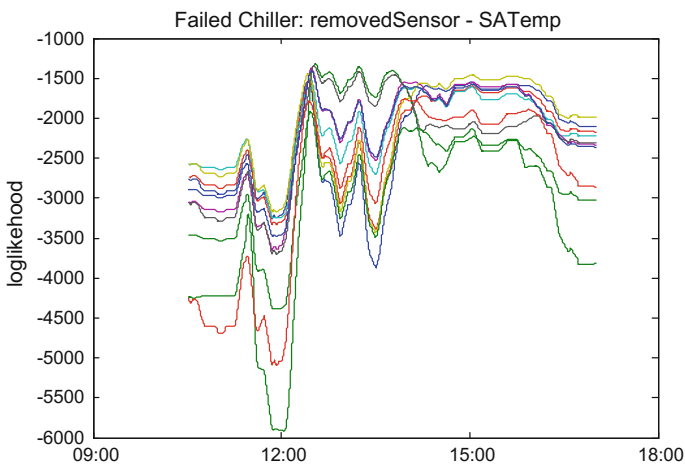

Fig. 5 A family of likelihood curves from a detected fault prior to clustering
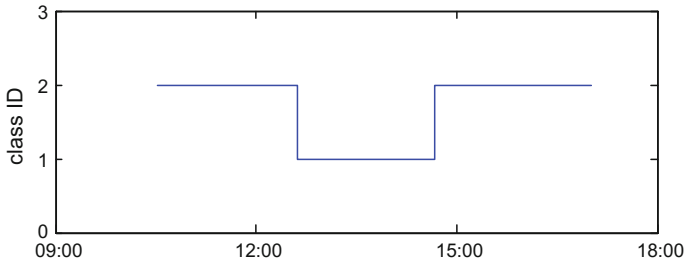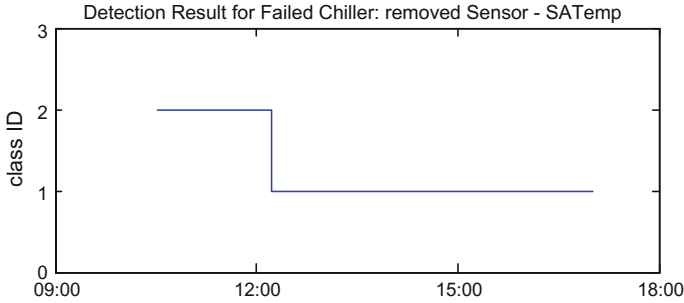
Fig. 6 Detection results for the fault after clustering and classification on the likelihood curves

If we manually remove one sensor feature from the optimized sensor combination, the fault would then become difficult to be detected. One example is shown in Figs. 7 and 8, for the same fault datasets—where the supply air temperature sensor data is not used for the HMM learning process. The likelihood curves drop to lower levels earlier that day, and the classified fault period is not correct. This example shows the importance of proper sensor features. Our approach can automatically select suitable sensor features rather than depend on domain expert, and assure successful detection results.

The above example proves that ERCE, as the data-driven sensor and feature selection algorithm, is crucial for the fault detection results. In fact, ERCE specifically performs well under varying seasonal conditions. The ERCE method selects features that are unique and relevant to the faults in the HVAC system. For different seasons, ERCE selects different combination of sensor data as the key features. Pattern for the winter dataset is different from the pattern for the summer dataset. The experimental results show that the sensor and feature selection process can ensure

Fig. 7 A family of likelihood curves without using optimal sensor combination

Fig. 8 Corresponding fault detection results

the success of the fault detection results for the HVAC system. It is proven that the ERCE method can improve the fault detection results, comparing to the results based on other feature selection clustering method, such as EAC k-means [18].

5   Conclusion and Discussion

This paper presents dynamic, machine-learning based techniques for automated sensor monitoring and fault detection in HVAC systems. This approach can be seen as a good combination of model-based methods and data-based methods. The main approaches are based on graphical modelling techniques such as HMMs, which encode probabilistic relationships among variables of interest. This approach is an ideal representation for combining prior knowledge and data. It does not need very detailed understanding of the physical system as in model-based approaches. It also does not need huge data sets as in the black-box approaches. Comparing with pure model-based or data-based approaches, it combines the strengths in both areas and can overcome their shortfalls by balancing the dependency on physical models and datasets.

One critical step to ensure the approach succeeds is that of the sensor feature selection process. This paper implements the ensemble rapid centroid estimation (ERCE) as the data-driven sensor and feature selection algorithm, which is the core method to ensure automated fault detection is achieved. Instead of choosing sensors manually, the ERCE method can automatically select representative features that are unique and relevant to the faults in the HVAC system. It also discards redundant sensors that are less crucial or have less system features. The experimental results show that this sensor and feature selection process can ensure the success of the fault detection results for HVAC systems. It is proven that the ERCE method can improve the fault detection results, comparing to the results based on other feature selection clustering methods, such as EAC k-means.

Planned future work includes comparing the selected sensor features from different buildings, and identifying the common features for multiple types of buildings.

By improving the generalizability of the sensor monitoring and fault detection approach, less modelling time can be saved, and more improvements in performance, such as fault detection accuracy, can be achieved.

References

1. M. Rolloos, HVAC systems and indoor air quality. Indoor Built Environ. **2**(4), 204–212 (1993)
2. A.G. Department of industry, innovation and science, in *Heating, Ventilation and Air-conditioning* (2016). http://industry.gov.au/Energy/EnergyEfficiency/Non-residentialBuildings/HVAC/Pages/default.aspx
3. K.W. Roth, J. Dieckmann, S.D. Hamilton, W. Goetzler, in *Energy Consumption Characteristics of Commercial Building HVAC Systems Volume III : Energy Savings Potential* (2002)
4. V. Vakiloroaya, B. Samali, A. Fakhar, K. Pishghadam, A review of different strategies for HVAC energy saving. Energy Convers. Manag. **77**, 738–754 (2014)
5. N. Gas, *Renewable Energy Annual 2002,*" vol. 0603 (2002)
6. J. Wall, Y. Guo, J. Li, S. West, A dynamic machine learning-based technique for automated fault detection in HVAC systems, in *2011 ASHRAE Annual Conference* (2011)
7. T.M. Rossi, J.E. Braun, A statistical, rule-based fault detection and diagnostic method for vapor compression air conditioners. HVAC&R Res. **3**(1), 19–37 (1997)
8. J. Schein, S. Bushby, N. Castro, J. House, A rule-based fault detection method for air handling units. Energy Build. **38**, 1485–1492 (2006)
9. P.P. Angelov, R.A. Buswell, V.I. Hanby, J.A. Wright, A methodology for modeling HVAC components using evolving fuzzy rules, in *2000 26th Annual Conference* of the *IEEE Industrial* Electronics *Society IECON 2000*, vol. 1, no. 1, (2000), pp. 247–252
10. Y. Song, Y. Akashi, J.-J. Yee, A development of easy-to-use tool for fault detection and diagnosis in building air-conditioning systems. Energy Build. **40**(2), 71–82 (2008)
11. B. Yu, D.H.C. Van Paassen, S. Riahy, General modeling for model-based FDD on building HVAC system. Simul. Pract. Theory **9**(6–8), 387–397 (2002)
12. J. Liang, R. Du, Model-based fault detection and diagnosis of HVAC systems using support vector machine method. Int. J. Refrig. **30**(6), 1104–1114 (2007)
13. H. Yoshida, S. Kumar, ARX and AFMM model-based on-line real-time data base diagnosis of sudden fault in AHU of VAV system. Energy Convers. Manag. **40**(11), 1191–1206 (1999)
14. D. Dehestani, S. Su, H. Nguyen, Y. Guo, J. Wall, F. Eftekhari, Comprehensive sensitivity analysis of heat ventilating and air conditioning (HVAC) system based on neural network model, in *10th International Conference on Healthy Buildings, Brisbane, Australia*, 8–12 July 2012, pp. 6–11
15. Y. Guo, J. Li, S. West, J. Wall, G. Platt, System and method for detecting and/or diagnosing faults in multi-variable systems, US20120185728 A1 (2012)
16. Y. Guo, J. Wall, J. Li, S. West, A machine learning approach for fault detection in multi-variable systems, in *The Tenth International Conference on Autonomous Agents and Multiagent Systems*, (2011), pp. 23–30
17. Y. Guo, D. Dehestani, J. Li, J. Wall, S. West, S. Su, Intelligent outlier detection for HVAC system fault detection, in *10th International Conference on Healthy Buildings, Brisbane, Australia,* 8–12 July, 2012, no. Vapnic (1995), p. 2
18. M. Yuwono, Y. Guo, J. Wall, J. Li, S. West, G. Platt, S.W. Su, Unsupervised feature selection using swarm intelligence and consensus clustering for automatic fault detection and diagnosis in Heating Ventilation and Air Conditioning systems. Appl. Soft Comput. **34**, 402–425 (2015)

19. Z. Yu, F. Haghighat, B.C.M. Fung, L. Zhou, A novel methodology for knowledge discovery through mining associations between building operational data. Energy Build. **47**, 430–440 (2012)
20. A. Kusiak, M. Li, F. Tang, Modeling and optimization of HVAC energy consumption. Appl. Energy **87**(10), 3092–3102 (2010)
21. N. Gaitani, C. Lehmann, M. Santamouris, G. Mihalakakou, P. Patargias, Using principal component and cluster analysis in the heating evaluation of the school building sector. Appl. Energy **87**(6), 2079–2086 (2010)
22. M. Yuwono, S.W. Sir, B.D. Moulton, Y. Guo, H.T. Nguyen, An algorithm for scalable clustering: ensemble rapid centroid estimation, in *Proceedings of the 2014 IEEE Congress on Evolutionary Computation, CEC 2014*, (2014), pp. 1250–1257
23. M. Yuwono, S.W. Su, B.D. Moulton, H.T. Nguyen, Data clustering using variants of rapid centroid estimation. IEEE Trans. Evol. Comput. **18**(3), 366–377 (2014)
24. M.A. Hearst, S.T. Dumais, E. Osman, J. Platt, B. Scholkopf, Support vector machines. IEEE Intell. Syst. **13**(4), 18–28 (1998)
25. J. Li, Y. Guo, J. Wall, S. West, Fault Detection for HVAC Systems, in *10th International Conference on Healthy Buildings, Brisbane, Australia,* 8–12 July 2012, pp. 12–15
26. D. Dehestani, F. Eftekhari, Y. Guo, S. Ling, S. Su, H. Nguyen, Online support vector machine application for model based fault detection and isolation of HVAC system. Int. J. Mach. Learn. Comput. **1**(1), 66–72 (2011)
27. S.R. West, Y. Guo, X.R. Wang, J. Wall, Automated fault detection and diagnosis of HVAC subsystems using statistical machine learning, in *12th International Conference of the International Building Performance Simulation Association* (2011)
28. S. Katipamula, M.R. Brambley, Methods for fault detection, diagnostics, and prognostics for building systems—A review, Part I. HVAC&R Res. **11**(1), 3–26 (2005)
29. EIA, International Energy Outlook 2009 (2009)