Synthesis and Refinement of Artificial HVAC Sensor Data for Supervised Learning in Data-Driven AFDD Techniques

David McCabe[1], Daniel Coakley[1], Catherine Conaghan[1] and Ruth Kerrigan[1]

1. ***Integrated Environmental Solutions Limited. (IES)***

Abstract

Up to 20% of the total energy used in developing countries is consumed within HVAC systems [1] with between 15-50% of this consumption being attributed to faulty operation [2,3,4]. Yu et al.[5] cite a UK survey conducted in 2000 which found prompt detection and diagnosis of HVAC faults can reduced the average plant consumption by more than 10-35%, similarly, in their analysis of VAV systems Lee 2010 [6] found considerable HVAC energy savings could be made through the adoption of Automated Fault Detection and Diagnostics (AFDD). Many approaches to HVAC AFDD have been developed [5], but the commercial viability of many of these techniques still needs to be thoroughly investigated. A history of known ‘tagged’ instances of faults within the data-set is essential for assessing and comparing the frequency of false alarms or detection success rates. It has been found in the authors experience that historical system data is often unavailable and inadequate for this purpose [6,7]. Therefore, simulation of faults using real test plant [8] or software [7] provides a promising alternative.

This paper proposes a scheme for the procurement and preparation of synthetic BEMS data in which faults are present using the IES-VE. This data is intended primarily for supervised learning in data-driven approaches but is also suitable for the testing of all HVAC AFDD methods. We identify and describe the characteristics that good quality training data should have, suggesting that fault observability should be considered when classifying (labelling) fault data. We suggest that the use of skewed data sets are permissible for the purposes of training data-driven classifiers and that brute-force simulation is a reasonable way of comprehensively sampling the typical operational envelope of many common systems. We conclude with a short discussion on our approach.

Introduction

AFDD has been underutilised in commercial HVAC applications [2,9], perhaps due to the required effort and perceived lack of returns in developing such a tool [6]. Bruton 2014 [4] states the use of AFDD will develop as a necessary tool for HVAC recommissioning citing the following reasons; 1) There are increasing financial and environmental pressures to reduce energy wastage, 2) More complex diagnostic tools are required to cope with increasingly sophisticated plant, 3) Emerging technologies such as wireless sensor networks make implementation evermore feasible. The Simulation Enhanced Integrated Systems for Model-based Intelligent Control(s) Project (EINSTEIN) is a European funded project which aims to develop and deploy a prototype building operation optimisation framework. A key aspect of EINSTEIN is to explore more sophisticated AFDD techniques than are currently used in existing BEMS however, since undertaking the project we have encountered issues in procuring historical building fault data for the testing of AFDD methods. In our experience BEMS data is rarely archived for long periods. Where faults do occur, these are not often logged appropriately to enable detailed tagging of operation data (e.g. date fault detected, date rectified). This paper outlines the philosophy and approach we have taken to procure labelled fault data for the training of data-driven AFDD techniques in EINSTEIN.

1. Fault Classification Scheme for HVAC Operation

Data-driven AFDD methods detect and identify faults at discrete time t using real valued data taken through the regular and synchronous sampling of a system's sensors $\mathbf{s}(t) = [s_0(t), s_1(t), ..]^T$. A set of n features $\mathbf{x}(t) = [x_0(t), x_1(t), ..x_n(t)]^T$ are formed from current sensor readings as well as preceding values if required $\mathbf{x}(t) = f(\boldsymbol{s}(t), \boldsymbol{s}(t-1), \boldsymbol{s}(t-2) \ldots)$ for input into a classifier. Fault identification is a classification problem where a classifier is used to map the feature space $\mathrm{g}: \mathbf{x} \mapsto \mathbf{y}$ into a finite number of prediction categories $y \in \{0,1,2...\}$ where $y = 0$ indicates a non-faulty operation prediction and $y = i$ indicates a prediction that fault type i has occurred. This is illustrated in (figure 1) using the notation $X_{y_{val}} \equiv \{\mathbf{x} \in X_U | g(\mathbf{x}) = y_{val}\}$ where X_U corresponds to the system's entire operational envelope. This categorisation topology of the feature-space is time independent. Fault detection can be viewed as the simpler binary classification problem where all of the fault prediction region are grouped into one $X_{fault} = X_1 \cup X_2 \cup ..$. An AFDD algorithm monitors a system's feature space trajectory, checking for incursions into the fault prediction regions by feeding sensor data through the classifier $(g \circ f): s \mapsto y$.

Classification ambiguity is undesirable therefore every point of the feature space should correspond exclusively to a single prediction category, that is to say no classification regions should overlap i.e. $\emptyset = \{X_i \cap X_j | i \neq j\}$. Faulty plant generally exhibits both faulty operation and non-faulty operation since faults may not always affect the system for example; the effects of the chiller breakdown are unlikely to be observed during periods when the system is in heating mode and the chiller is non-operational. Fault prediction regions should therefore only correspond to detectable faulty operation which is distinguishable from baseline operation. Similarly two different fault types might sometimes produce the same symptoms making them indistinguishable in feature space. Regions of feature space in which multiple fault regions overlap can be considered as distinct fault regions in their own right, as would any regions corresponding to multiple faults acting simultaneously.

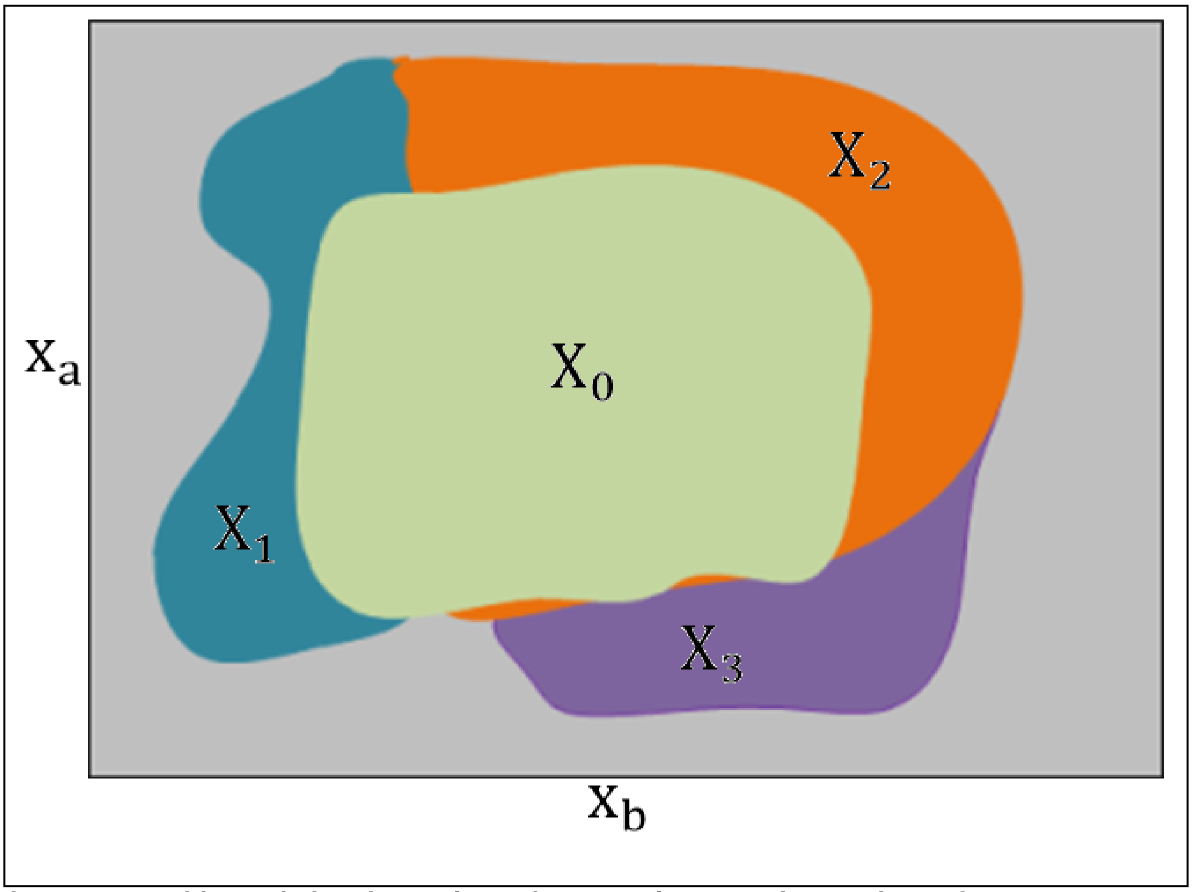

Figure 1: Graphic Depicting the Faulty and Non-Faulty Operation Regions of Feature Space

2. Desirable Characteristics of Labelled Fault Data

The different types of data-driven classifier (e.g. Artificial Neural Networks, Support Vector Machines, Random Forests etc...) are essentially parameterised functions capable of representing the prediction boundaries in $\mathbb{R}^n$ feature-space. A data-driven classifier's prediction boundaries are optimised through supervised learning, where the classifier's parameter values are adjusted to minimise the classifier's prediction error on a training dataset D_{train} in which each data point $\mathbf{x}_i$ is labelled with the correct prediction category y_i.

$$D_{train} = \{(\mathbf{x}_i, y_i) | \mathbf{x}_i \in \mathbb{R}^n, y_i \in \{0,1,2 \ldots\}\}_{i=1}^{N_{epoch}}$$

A classifier's performance is highly dependent on the quality of the data used to train it. There are a number of characteristics good quality training data should possess; first and foremost the labelling of the data should be accurate preferably using fault labels corresponding to detectable faults only as outlined in the former section. Exhaustive sampling of the feature space particularly around the boundary regions can ensure that class boundaries are positioned accurately with little margin for uncertainty, furthermore an even spread or distribution of points insures that the classifier accuracy is even across the feature space and not bias towards a particularly dense cluster of points. Training data should have diversity or spread of data and even distribution, however these are difficult to obtain if the dimension of the feature space is high and if the data procurement method suffers from sampling bias. We anticipate the high dimensionality issue is alleviated for fault detection in most systems for two reasons; many features are typically correlated and therefore data points lie within or around a much lower dimensional subspace in the feature space (lower than that found using principle component analysis which only considers linear correlations), secondly, most of a system's operational envelope corresponds to atypical system behavior which is unlikely to ever be encountered, comprehensive sampling can therefore be limited to typical or reasonable operating region.

3. The Procurement of Raw Fault Data using the IES-VE

Data-driven AFDD techniques can be used with proactive testing [8], this involves intentionally seeding adverse system problems in real plant while training data is collected however, this may be time consuming and prohibit since in could result in the damage of equipment, simulation is therefore the only method for the procurement of training data in many cases [7]. We have developed a utility which produces large volumes of artificial sensor data from AHU simulations subjected to common operational faults using the IES-VE software package.

The IES-VE software suite allows users to import or construct a geometrical model of the client building on a 3D canvas using the ModelIT module. All time varying simulation inputs are defined as profiles and added to the profile database within the ApPro module (figure 2). An alternative module ApHVAC is provided for the modelling of HVAC plant (figure 3). Users can define air and waterside loop configurations assigning the profiles which drive the plant serving them. The IES-VE was originally intended as a building design tool simulating idealised control for compliance, but it is now also being used for the simulation of building operation for instance, freeform profiles permits the import of actual plant sensor data to drive realistic calibrated building models. We have used IES-VE to simulate both normal and faulty operation of a number of typical AHU configurations.

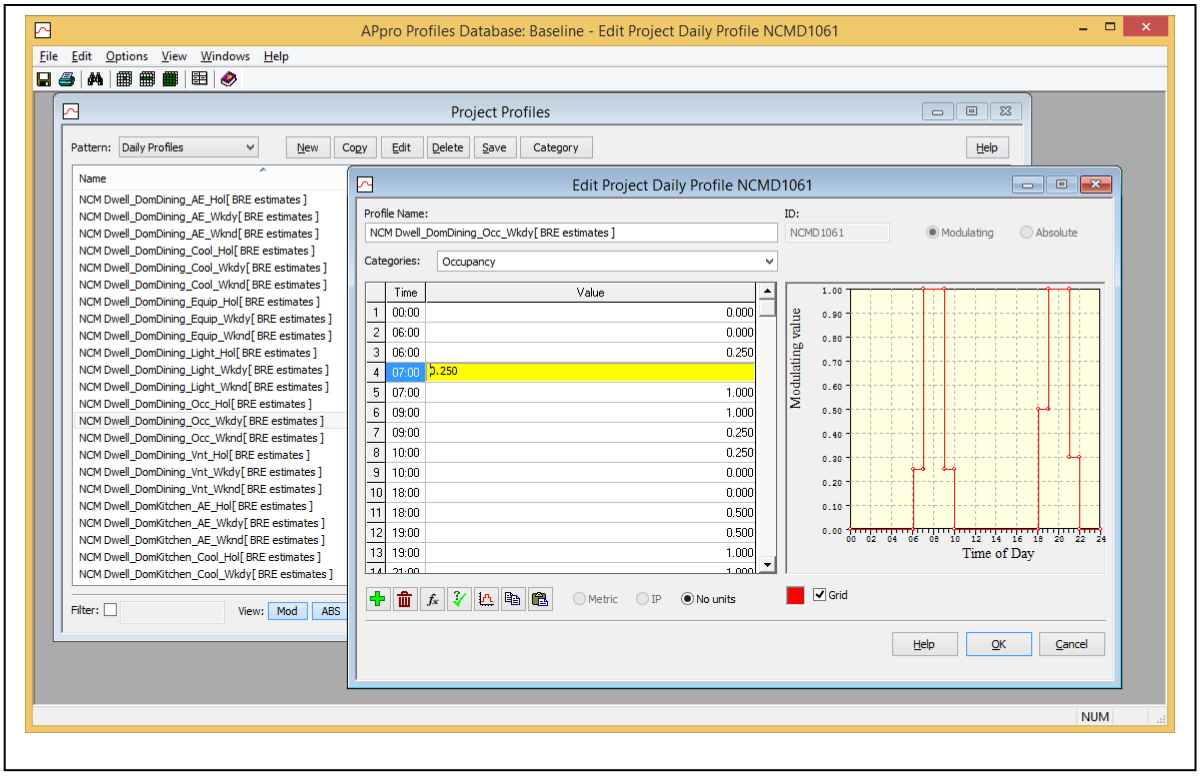

Figure 2: Profile Editing using IES-VE's ApPro

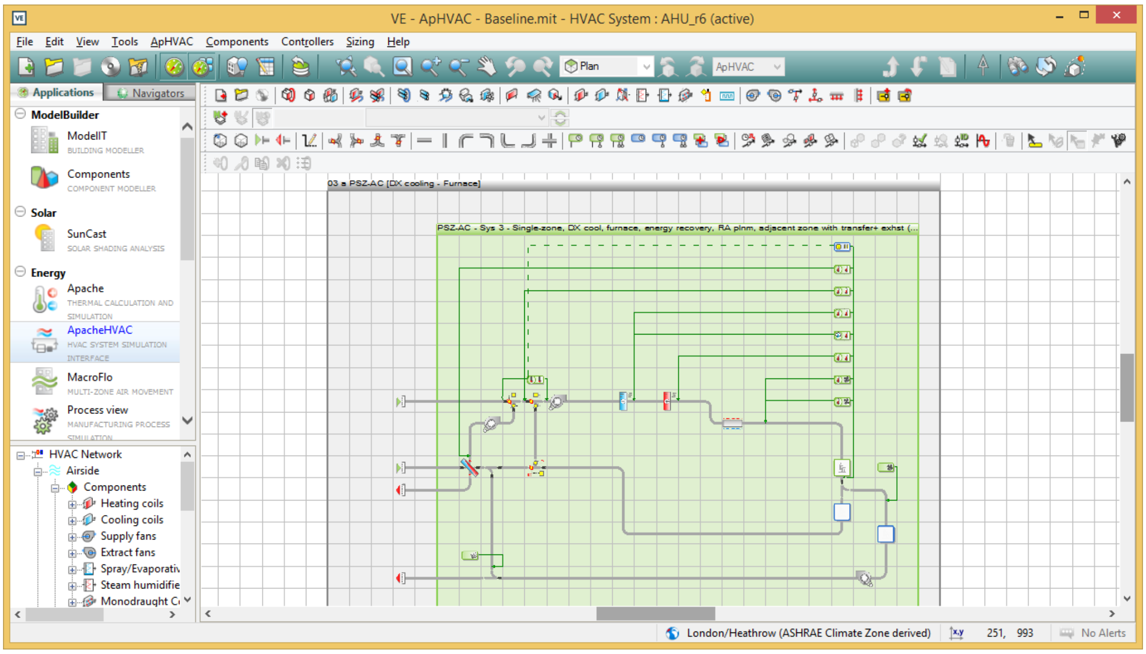

Figure 3: AHU editing in IES-VE's ApHVAC

3.1. Simulating Building Operation using IES-VE

Standard IES-VE models simulate HVAC plant functioning as intended, these can be used for the procurement of sensor data representing the non-faulty operation of equipment. A baseline model of a building and plant is first set up, this simulates arbitrary operation of the system over a three month period with results output at 6 minute reporting intervals. The baseline model is then cloned, producing multiple variant models in which the operational input parameters (profiles) are randomly perturbed to produce a set of results files corresponding to system operation over a wide range of the system's non-faulty envelope. The raw sensor data is extracted from the results files and added to a database for storage. The variant models can be generated and simulated repeatedly to produce a very large and diverse dataset.

3.2. Simulating Faulty Building Operation using IES-VE

Little in terms of research has been conducted in the area of modelling strategies for common HVAC faults using building design software [7,10]. The IES-VE suite does not cater for the explicit modelling of equipment faults however, its robust first principles core engine (as opposed to more data driven tools), permits implicit fault simulation using the two strategies presented in [7]; firstly extending the idealised model to make it capable of simulating adverse system conditions and secondly the perturbing numerical values at simulation runtime via free form profile manipulation to mimic malfunctioning equipment. These techniques are applied to produce a set of baseline models corresponding to faulty operation from which sensor data can be procured in the same way section 3.1 describes. A new baseline model is required to simulate every individual type of system fault which may occur as well as every combination of faults which can occur. This is extremely time intensive and so far we have limited ourselves to the detection of the most common faults in the systems we have studied.

4. The Refinement of Fault Data

Post processing of the raw simulated sensor data taken from the IES-VE simulations is required to put it into the form suitable for data-driven fault classifiers described in section 2. Sections 4.1 and 4.2 present a simple approached for the accurate labelling and balancing of data respectively. This approach may also be used to put data procured from any source into the form we desire provide that; the presence of all faults in the data are known and the data set represents a comprehensive coverage the system's operational feature space envelope.

4.1. Labelling of Fault Data

The volume occupied by a cloud of data points in a multi-dimensional feature space can be found using the technique of kernel density estimation (KDE). A Gaussian kernel is placed over each data point so that the sum of the kernels correspond to a normalised density distribution of the dataset. Next a tolerance is chosen between zero and the peak height of the selected kernel, every point of the feature space which produces a larger value when sampled in the density distribution function is considered to be inside the region of data points. We apply this technique to detect and address any overlap between the different classes of points.

Initially each simulation data point is assigned a label based on the fault or faults which were present in the simulation from which it was produced. All data samples in which no faults are present can be immediately labelled as belonging to the non-faulty operation region of feature space. Next we consider the detectability in all of the data in which faulty operation is present discarding and data points in which a fault is present but undetectable. KDE is used to identify the volume of the feature space which corresponds to non-faulty operation X_0. Any data points procured from faulty plant simulations which are found to reside in the non-faulty operation region of feature space are discarded.

Similarly, we can iterate over the fault classes and check for overlap between them. When overlap between two fault categories is identified, all points in the overlapping region are put into a new category corresponding to a prediction region in which the classifier is uncertain of which individual fault is present. This may be undesirable and indicates that different features should be selected to make faults more resolvable.

4.2. Balancing of Fault Data

Data-driven fault classifiers should be ideally be trained using data spread evenly across the typical operational envelope of the system. This is typically not a characteristic of the data procured from the variant simulations and consequently, we discard (under-sample) a proportion of data points located in the denser sampling regions of the feature space. Kernel density estimation can again be used to determine which data points to discard after the fault labelling process or alternatively this can be done when forming the density functions during the labelling procedure. The later approach works as follows, before adding a new kernel to the density function we sample the current semi-constructed density function at the location of the new data point (corresponding to the new kernel). If the semi-constructed density function returns a value which is too large - for instance the height of two kernels, we discard the data point and do not add the new kernel to our density function.

Concluding Remarks

A computational approach to fault modelling like the one we have outlined in this paper shows promise in being able to produce large volumes of reliable data quickly. This can be used to compare the detection rate and frequency of false alarms in different classifiers, allowing progression of the AFDD aspect of EINSTEIN. Further study is needed in evaluating the use of this data for the training of data-driven AFDD classifiers which are to be deployed on real buildings, since we acknowledge there is typically a large disparity between data procured from real plant and simulated data.

This paper proposes that data-driven fault classifiers should be trained using data spread evenly across the typical operational envelope of the system. This results in class imbalance when the $\mathbb{R}^n$ dimensional volumes of feature space corresponding to the different fault categories vary dramatically in size. We note this may contravene some widely reported guidelines which states this can lead to poor performance in data-driven classifiers however, we share the view of (Batista et al. 2004 [11]) who find skewed classes may be permissible when there is no overlap between the different classes of points - something we addressed by the design of our labelling scheme.

Our approach samples the classification of a feature space through simulating the time evolution of HVAC plant, in effect we are sampling along numerous paths or trajectories through the space. Considering a HVAC system as a group of attractors in feature space, it is conceivable that this approach inevitably leads to extreme sampling bias, something we have explicitly stated we would like to avoid. We can offer two explanations for this; (i) the most sampled regions will correspond to the system's typical operational envelope which is the region of interest, (ii) hysteresis effects and the slower nature of building dynamics dictated that it is necessary to simulate (precondition) for a long period prior to sampling.

Acknowledgments

The EINSTEIN project acknowledges the support received from the European Union's FP7 program through the Marie Curie Industry-Academia Partnerships and Pathways (IAPP) program, under grant agreement no: 611012

References

[1] L. Pérez-Lombard, J. Ortiz, and C. Pout, “A review on buildings energy consumption information,” Energy and Buildings, vol. 40, no. 3, pp. 394–398, Jan. 2008.

[2] S. Katipamula and M. R. Brambley, “Methods for fault detection, diagnostics, and prognostics for building systems - A review, Part I,” HVAC and R Research, vol. 11, no. 1, pp. 3–25, 2005.

[3] X. Li and J. Wen, “Review of building energy modelling for control and operation,” Renewable and Sustainable Energy Reviews, vol. 37, pp. 517–537, Sep. 2014.

[4] K. Bruton, P. Raftery, B. Kennedy, M. M. Keane, and D. T. J. O'Sullivan, "Review of automated fault detection and diagnostic tools in air handling units," Energy Efficiency, vol. 7, no. 2, pp. 335–351, 2014.

[5] Y. Yu, D. Woradechjumroen, and D. Yu, "A review of fault detection and diagnosis methodologies on air-handling units," Energy and Buildings, vol. 82, pp. 550–562, Oct. 2014.

[6] S. Lee and F. Yik, "A study on the energy penalty of various air-side system faults in buildings," Energy and Buildings, vol. 42, no. 1, pp. 2–10, Jan. 2010.

[7] P. Haves, "Fault modelling in component-based HVAC simulation," Proceedings of Building Simulation'97, 1997.

[8] M. Najafi, "Fault Detection and Diagnosis in Building HVAC Systems," Jan. 2010.

[9] S. Katipamula and M. Brambley, "Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part II," HVAC&R Research, vol. 11, no. 2, pp. 169–187, 2005.

[10] Mangesh Basarkar, X. Pang, L. Wang, P. Haves, and T. Hong, "Modeling and simulation of HVAC faults in EnergyPlus," IBPSA Building Simulation, pp. 14–16, 2013.

[11] G. Batista, R. Prati, and M. Monard, "A study of the behavior of several methods for balancing machine learning training data," ACM Sigkdd Explorations Newsletter, vol. 6, no. 1, pp. 20–29, 2004.