Energy and Buildings 86 (2015) 698–707

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Robust model-based fault diagnosis for air handling units

Timothy Mulumba[a], Afshin Afshari[a,*], Ke Yan[a], Wen Shen[a], Leslie K. Norford[b]

[a] Department of Engineering Systems and Management, Masdar Institute of Science and Technology, PO Box 54224, Abu Dhabi, United Arab Emirates
[b] Department of Architecture, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

ARTICLE INFO

Article history:
Received 10 August 2014
Received in revised form 1 October 2014
Accepted 29 October 2014
Available online 4 November 2014

Keywords:
Fault detection and diagnosis
Air handling units
ARX model
Machine learning techniques
Parameter estimation
Recursive least squares

ABSTRACT

Fault detection and diagnosis (FDD) for heating, ventilation and air conditioning (HVAC) equipment significantly impact energy consumption in both residential and commercial buildings. In most modern building management systems (BMS), HVAC historical data logs in large quantity and high resolution are recorded and available for further online or offline analysis, including automated FDD. In this paper, a model-based fault diagnosis method is developed by applying support vector machine (SVM) techniques to model parameters recursively calculated by an online estimator. The estimator presumes an autoregressive time series model with exogenous variables (ARX). A real-world air handling unit (AHU) dataset containing process variables measured at regular intervals is pre-processed by the online estimation algorithm. Each data vector in the original dataset (measured), together with a small number of appropriately selected lags, is converted to a parameter vector representing the state at the same instant. The set of parameter vectors is sub-divided into classes by SVM, enabling fault classification. Validation via experimental data demonstrates that the proposed hybrid approach produce superior performance measured by *F*-measure scores compared to alternative methods. Robustness to model uncertainty is also established.

1. Introduction

Poorly maintained HVAC equipment represents up to 30% of total energy consumed in commercial buildings [1]. By improving control and maintenance, FDD techniques significantly impact HVAC energy efficiency. Despite the integration of increasingly sophisticated monitoring and control functions in modern BMS to meet the growing requirement for reliable and economic operation, these systems are usually not able to detect non-catastrophic HVAC malfunction before the scheduled maintenance nor do most malfunctions impact end-user thermal comfort to any noticeable degree. Meanwhile, building management systems monitor various sensors and store the data in large databases, which make it possible to save energy by mining the data and scheduling necessary preventive maintenance. In hot and humid climates, the potential impact of this function is compounded by the fact that the air-conditioning load may even represent up to 60% of the total grid electricity consumption. Detection methods based on energy/cost benchmarking of HVAC equipment can detect faults but they are best suited to offline (after the fact) analysis and do not account for weather/schedule variability. Accuracy and detection speed are important from an operational point of view and require the use of high frequency process measurements.

The AHU is an important component of HVAC plants. Its main function is to control indoor air quality by continuously supplying conditioned outdoor air. The mixed air, carefully dosed mixture of outdoor air and return air from the indoor zone, is conditioned in the AHU. The conditioning is achieved by passing the mixed air flow over the heating and/or cooling coils, as necessary. Once the air has been conditioned to the prescribed temperature/humidity levels, it is supplied to the indoor environment. The prototypical AHU system considered in this study does not include an energy recovery system.

Categories of faults in AHU include mechanical failures, control problems, design errors and inappropriate operator intervention. A series of papers summarize up to 15 major faults for an AHU system including exhaust air damper stuck, return fan fault, cooling coil valve control fault, outside air damper stuck, cooling coil valve stuck, heating coil valve leakage, outside air damper leakage, heating coil fouling, heating coil capacity reduction, AHU duct leakage, outside air temperature sensor bias, air filter area blockage, mixed air damper unstable, supply fan control unstable and outside air damper stuck [2–5]. In this work, we investigate faults most likely to occur in summer operating conditions. Real world data provided

* Corresponding author. Tel.: +971 2 8109149; fax: +971 2 8109901.
E-mail address: aafshari@masdar.ac.ae (A. Afshari).

http://dx.doi.org/10.1016/j.enbuild.2014.10.069
0378-7788/

Nomenclature

E_{EX}	enthalpy rate of exhaust air (kW)
E_{MA}	enthalpy rate of mixed air (kW)
E_{OA}	enthalpy rate of outside air (kW)
E_{RA}	enthalpy rate of return air (kW)
E_{SA}	enthalpy rate of supply air (kW)
F	*F*-measure
FN	false negative
FP	false positive
$G(t)$	gain function
H_{MA}	mixed air humidity ratio (kg/kg)
H_{OA}	outside air humidity ratio (kg/kg)
H_{RA}	return air humidity ratio (kg/kg)
K	number of faults
M	number of measured process variables
$\dot{m}$	mass flow rate (kg/s)
$\dot{m}_{MA}$	mass flow rate of mixed air (kg/s)
$\dot{m}_{OA}$	mass flow rate of outside air (kg/s)
$\dot{m}_{SA}$	mass flow rate of supply air (kg/s)
N	number of data vectors
$P(t)$	variance-covariance matrix
Q_{coil}	heat transfer rate of the cooling coil (kW)
t	time step
T	temperature (°F)
TP	true positive
$v_M(t)$	measured process variable
V	objective function
X	training dataset
$x(t)$	data vector
$y(t)$	observed stationary process
$z(t)$	class label
Z	set of class labels
Z_1	class labels training subset
Z_2	class labels testing subset
Δ	sampling interval
$\epsilon(t)$	prediction error
$\theta(t)$	parameter vector
Θ	set of parameters
Θ_1	parameter vector training subset
Θ_2	parameter vector testing subset
λ	forgetting factor
τ_f	exponential forgetting time constant
$\phi(t)$	vector of exogenous variables

directly to us by the authors of ASHRAE report 1312-RP [6–8] is utilized to train and validate our algorithm.

The proposed FDD algorithm implements a novel hybrid method combining an autoregressive model with exogenous inputs (ARX) and support vector machine (SVM) technique to detect and diagnose AHU faults. First, the original dataset is divided into two parts; the first part of data with known fault types constitutes the training dataset. The second part of data, where fault labels are initially hidden, is called the testing dataset and is used for validation purposes. Both training and testing subsets are pre-processed by an online model identification algorithm and converted into a set of parameter vectors, one per sampling instant. It is essential to retain the sequential disposition of the data at this stage, given the assumed time series nature of the process. Thereafter, the parameter dataset is randomly shuffled and subdivided into training (2/3 of the data) and testing (1/3 of the data) subsets. The random shuffling simulates a stationarized process. The resulting training and testing parameter subsets are then processed by SVM first for supervised learning, second for validation. In the validation stage, *F-measure* scores are determined (see details in Section 3.4). High *F*-measure is achieved indicating superior prediction accuracy and exceptionally low false alarm rates in identifying different types of AHU faults. Alternative methods are tested and shown to be inferior. The proposed method is shown to be fairly robust to model parameter uncertainty.

1.1. Related work

Studies on FDD are numerous and various methods have been reported for AHUs in the literature. Venkatasubramanian et al. [9–11] present an overview of FDD methods applicable to building HVAC systems; popular methods for HVAC systems which process large databases include expert systems, neural network models, principal component analysis (PCA), support vector machines (SVM), and combinations of these techniques.

Henley [12] first applied the expert system to diagnose faults. Henley's inference algorithm used a set of if-then rules to determine the state of a system and provide a possible source of the fault. Ramesh et al. [13] introduced a hierarchical classification to improve Henley's approach by focusing on the information processing tasks underlying diagnostic reasoning. Their framework was successfully applied to the development of knowledge-based diagnostic systems for several processes. Schein et al. [14] developed a rule-based fault detection method derived from mass and energy balances for AHUs. Their rule-based system was computationally simple enough that it could be embedded in commercial building automation and control systems. Norford et al. [15] developed two methods for FDD in HVAC equipment. One of the methods used first-principles-based models of system components. The data used by this approach was obtained from sensors typically installed for control purposes. The second method was based on semi-empirical correlations of sub-metered electrical power with flow rates or process control signals. As a result, the first-principles-based models require a larger number of sensors than the electrical power correlation models, although the latter method requires power meters that are not typically installed. Li and Braun [16] formulated model-based FDD for vapor-compression air conditioning equipment in a generic way by developing a physical decoupling methodology as an alternative to mathematical decoupling.

Farell and Roat [17] modified the traditional back-propagation neural network for fault diagnosis by incorporating a small amount of process knowledge which minimized data limitations (since neural networks perform only as robustly as the data from which they are trained) and maximized the neural network's performance. Lee et al. [18] developed a general regression neural network (GRNN) model at the subsystem level of AHU. Their results showed that the GRNN model is an accurate and reliable estimator for highly nonlinear and complex AHU processes. Magoulès et al. [19] developed a recursive deterministic perceptron (RDP) neural network for FDD at the building level. Based on their results which showed a higher than 97% generalization performance, they proposed a new diagnostic architecture that reported both the source of the faults and their degradation likelihood.

Wang et al. [20,21] and Xiao [22] applied the PCA method and diagnosed sensor faults for both chiller and AHU systems. They developed two models based on PCA analysis of sensor readings which could group sets of correlated variables and capture the systematic trends of the system. The models were also used to estimate the bias magnitudes of the sensors. Du et al. [23] applied a PCA model to detect sensor faults including fixed bias, drifting bias and complete failure in air dampers and VAV terminals.

In order to maximize the inherent advantages of each technique, many researchers use various combinations of the aforementioned techniques in the last two decades; Becraft and Lee [24] combined neural networks and expert systems to develop a diagnostic