Applied Energy 188 (2017) 200–214

Contents lists available at ScienceDirect

Applied Energy

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

Fault detection and diagnosis of chillers using Bayesian network merged distance rejection and multi-source non-sensor information

Zhanwei Wang [a], Zhiwei Wang [a,*], Suowei He [a], Xiaowei Gu [a], Zeng Feng Yan [b]

[a] School of Environmental and Municipal Engineering, Xi'an University of Architecture and Technology, 13 Yanta Road, Xi'an 710055, PR China
[b] School of Architecture, Xi'an University of Architecture and Technology, 13 Yanta Road, Xi'an 710055, PR China

HIGHLIGHTS

- Fuse distance rejection (DR)/multi-source information (MI) into Bayesian network (BN).
- A novel chiller FDD method based on DR-MI-BN is developed effectively.
- Identify new types of faults and update fault library by merging DR into BN.
- False alarm rate can meet different users' requirements by adjusting the DR.
- FDD performance are improved significantly by merging MI into BN.

ARTICLE INFO

Article history:
Received 28 July 2016
Received in revised form 12 November 2016
Accepted 30 November 2016

Keywords:
Bayesian network
Chiller
Distance rejection
Fault detection and diagnosis
Multi-source non-sensor information

ABSTRACT

Applying the fault detection and diagnosis (FDD) techniques to chillers is beneficial to reduce building energy consumption and to enhance the energy efficiency of buildings. The purpose of this study is to propose a chiller FDD method with better performance for field implementation. The technological paths are as follows: (i) in order to identify new types of faults and to update the FDD fault libraries, a distance rejection (DR) technique is merged into the Bayesian network (BN) by transforming the chiller FDD problem into a single-class classification problem. Furthermore, the DR can be tuned to obtain an adjustable false alarm rate (FAR); (ii) to increase the diagnostic accuracies of known (or existing) faults and the identification accuracies of new types of faults, multi-source non-sensor information (MI) is merged into the BN, i.e., maintenance records and repair service history, healthy states of related equipment and on-site observed information. A novel chiller FDD method based on BN merged DR and MI (DR-MI-BN) is proposed in this study. The performance of this proposed method is evaluated by using the experimental data from ASHRAE RP-1043. Test results show that the FAR can be tuned for different users' requirements, and that merging the MI significantly improves the diagnostic accuracies of known faults from 77.2% to 99.8% at most (for refrigerant leakage) and the identification accuracies of new types of faults from 56.6% to 99.6% at most (for NF7).

1. Introduction

The building sector is widely recognized as a major consumer of both energy and resources, which currently takes up approximately 41% of the total primary energy in the United States, 40% in the European Union, and 20% in China [1–3]. Chiller plants account for more than 40% of the total energy used in commercial and industrial buildings [4] and have a major impact on comfort conditions and building maintenance costs. Energy usage of a chiller can be increased due to component faults: refrigerant leakage, condenser fouling, and so on. Much of the waste caused by faults could be prevented with widespread adoption of automated fault detection and diagnosis (FDD) [5]. Applying the FDD techniques to chillers is beneficial to reduce building energy consumption and to enhance the energy efficiency of buildings.

Over the past two decades, the topic of FDD in HVAC&R (Heating, Ventilating, Air-Conditioning and Refrigeration) has been an active area of research [6–10], and many researchers have devoted themselves to developing FDD methods for chillers. A comprehensive overview in the context of building-specific application classifies FDD methods into three categories [11]: quantitative model-based methods, qualitative model-based methods, and data-based methods. In spite of the progress and effort made,

* Corresponding author.
E-mail address: wzhiwei-atu@163.com (Z. Wang).

http://dx.doi.org/10.1016/j.apenergy.2016.11.130
0306-2619/

Nomenclature

c_0	coefficient representing the distance rejection of normal
c_i	coefficient representing the distance rejection of fault F_i
CL	T^2 control limit
FAR	false alarm rate
F_i	ith known fault
HS	healthy states of the related equipment
MR	maintenance services on chiller
m	dimension of the multivariate $\mathbf{X}$
n	sample number of the considered class
NF	new type of fault
P	probability value
pdf	probability density function
SI	on-site observed information
$\mathbf{x}$	multivariate observation
$\mathbf{X}$	multi-dimensional variable

Greek symbols

α	significance level
$\boldsymbol{\mu}_0$	mean vector of multivariate $\mathbf{X}$ belonging to normal
$\sum_0$	covariance matrix of multivariate $\mathbf{X}$ belonging to normal
$\boldsymbol{\mu}_i$	mean vector of multivariate $\mathbf{X}$ belonging to fault F_i
$\sum_i$	covariance matrix of multivariate $\mathbf{X}$ belonging to fault F_i

Subscript

q	number of node SI

widespread adoption in the field still is slow. Inadequate fault data and the impact of false alarm rate (FAR) of an FDD method on a user are two of main reasons.

Inadequate fault data is one of the key issues for field implementation of FDD methods for chillers. Although the most common faults with highest energy and operational impacts have been well documented for chillers [9,12], it is hard to obtain the full-set data including all common faults in the field. According to the survey in [12], there are seven common chiller faults. The equipment manufacturers and building owners are reluctant to perform such experiments as manually introducing faults to their equipment and systems because of the high costs. With the development of technology, chillers become more and more complex, which causes more types of faults. It is unrealistic to obtain a full-set of fault data in the field. However, fault data for one or two common faults are usually available in the field. Once a special fault occurs, its data will be stored in the management system. An effective solution, with respect to inadequate fault data, is the ability of identifying new types of faults (faults not included in training datasets) and updating the FDD fault libraries.

For fault detection, FAR is also of great significance for FDD systems in field applications. Different users have different tolerances for FAR. For critical systems (i.e., those related to life safety or having enormous costs), users require FDD methods applied to these systems to be very sensitive to fault detection so as not to miss any instance of faults. Therefore, these sensitive methods usually generate higher FAR. For noncritical applications (i.e., those not related to life safety or having enormous costs), the FDD methods should minimize the number of false alarms. If an FDD system generates a large number of false alarms, the users may become frustrated, lose confidence in the system, and even disable it completely in response [11]. Therefore, a better way is having an adjustable FAR that can satisfy different requirements.

In addition to the two aforementioned main reasons, there are also three additional aspects needed to be considered: (i) generally, variables measured on chillers are highly non-linear; (ii) because of limits to the measured data quantity, data acquisition accuracy and uncertainty of mapping relations between fault symptoms and faults, various uncertainties exist in the FDD process; (iii) a variety of information resources besides sensor data (multi-source non-sensor information) exist for chillers, such as maintenance records and repair service history, healthy states of related equipment and on-site observed information. For instance, the refrigerant overcharge fault seldom occurs suddenly if refrigerant charge service is absent. The chiller is at a higher risk of refrigerant leakage if grease stains are observed on pipe joints or interfaces.

Therefore, the desired attributes of chiller FDD methods for field implementation include:

- high detection and diagnosis accuracies,
- the ability to adjust its FAR,
- the ability to identify new types of faults and to update FDD fault library,
- the ability to address non-linear variables and uncertainties, and to merge multi-source non-sensor information.

A literature survey shows that various FDD methods for chillers have been developed, such as quantitative model-based methods [13–15] and rule-based methods [16–19]. However, for large systems (many inputs, outputs and states, e.g., chillers), detailed models developed on the basis of detailed information about fault mechanism are complex and time consuming. These rule-based methods, to a large extent, depend on the expertise and knowledge of developers [11], and generally have low accuracies, especially at slight severity levels. The data-based methods for chiller FDD have caught much research attention [20–27]. Principal component analysis (PCA) is one of the popular data-based methods, and has already been applied to AHU sensor fault diagnosis [21], VAV system sensor fault diagnosis [22], chiller sensor FDD [23]. However, PCA is not efficient for chiller FDD because it is a linear method [24]. Han et al. [25,26] applied a support vector machine (SVM) to detect and diagnose chiller faults, in which the FDD problem was viewed as a multi-class classification problem. Zhao et al. [24,27] introduced a support vector data description (SVDD) for chiller FDD, in which the FDD problem was considered to be a single-class classification problem. The test results showed that it was more powerful than multi-class classification. Single-class classification has the ability to identify new types of faults and the potential to adjust the FAR. Yet the SVDD-based FDD method has no ability to address uncertainties, which is one of the desired attributes of chiller FDD methods for field implementation, and its FDD results are Booleans (i.e., yes or no), which means that an observation belonging to only one of the known faults (faults included in training datasets) may be simultaneously diagnosed as more than one fault. For instance, in Zhao et al. [27], 32 percent of the total samples belonging to refrigerant leakage were simultaneously diagnosed as refrigerant leakage and refrigerant overcharge. Such results will confuse most users, and diagnostic efficiency is reduced by checking all alarmed faults. Therefore, a new solution is needed to realize the single-class classification.

A Bayesian network is a probabilistic graphical model, and is considered to be one of the most useful models in the field of prob-