available at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/ijrefrig

Important sensors for chiller fault detection and diagnosis (FDD) from the perspective of feature selection and machine learning

H. Han [a,*], B. Gu [a], T. Wang [a], Z.R. Li [b]

[a] Institute of Refrigeration and Cryogenics, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai 200240, PR China
[b] Institute of HVAC & G, School of Mechanical Engineering, Tongji University, 1239 Siping Road, Shanghai 200092, PR China

ARTICLE INFO

Article history:
Received 7 June 2010
Received in revised form 23 July 2010
Accepted 17 August 2010
Available online 21 August 2010

Keywords:
Sensor
Chiller
Compression system
Detection
Genetic
Fault

ABSTRACT

The benefits of applying automated fault detection and diagnosis (AFDD) to chillers include less expensive repairs, timely maintenance, and shorter downtimes. This study employs feature selection (FS) techniques, such as mutual-information-based filter and genetic-algorithm-based wrapper, to help search for the important sensors in data driven chiller FDD applications, so as to improve FDD performance while saving initial sensor cost. The 'one-against-one' multi-class support vector machine (SVM) is adopted as a FDD tool. The results show that the eight features/sensors, centered around the core refrigeration cycle and selected by the GA-SVM wrapper from the original 64 features, outperform the other three feature subsets by the GA-LDA (linear discriminant analysis) wrapper, with an overall classification correct rate (CR) as high as 99.53% for the 4000 test samples randomly covering the normal and seven typical faulty modes. The CRs for the four cases with FS are all higher than that without FS (97.45%) and the test time is much less, about 28–36%. The FDD performance for normal or each of the faulty modes is also evaluated in details in terms of hit rate (HR) and false alarm rate (FAR).

Capteurs importants pour la détection et le diagnostic d'anomalies des refroidisseurs du point de vue des choix des caractéristiques et des connaissances du système

Mots clés : Capteur ; Refroidisseur ; Système à compression ; Détection ; Génétique ; Anomalie

1. Introduction

Automated fault detection and diagnosis (AFDD) along with prognostics is the cornerstone for automated condition-based maintenance, whose wide-spread adoption will help cut down much of the waste caused by poorly maintained, degraded and/or improperly controlled equipment. Although there is a wealth of literature related to AFDD for critical processes,

* *Corresponding author*. Tel./fax: +86 21 3420 6260.
E-mail address: happier_han@126.com (H. Han).
0140-7007/$ – see front matter
doi:10.1016/j.ijrefrig.2010.08.011

Nomenclature

b	bias or threshold for the discriminant function
c	target class vector
C	penalty constant (also called *slack penalty*)
ConfMat	confusion matrix
CR	correct rate
FAR	false alarm rate
FN	number of false negative samples
FP	number of false positive samples
FWC	flow rate of condenser water
HR	hit rate
M_j	mean of the jth variable/feature
MCR	misclassification rate
p	matrix of the posterior probabilities for the classifier
PO_feed	pressure of oil feed
S_a	training set of two class
S_x	possible values' set for X
S_y	possible values' set for Y
TCO	temperature of condense water out by RTD
TEO	temperature of evaporator water out
TN	number of true negative samples
TP	number of true positive samples
TRC	saturated refrigerant temperature in condenser
TR_dis	refrigerant discharge temperature
TWCO	temperature of condense water out by thermistor
TWI	temperature of city water in
VE	position of the electronic valve installed in the evaporator water loop
w	weight vector for SVM
X	discrete random variable
$\hat{X}$	normalized sample matrix
Y	discrete random variable
Z	observed sample matrix

Greek symbols

α_i	Lagrange multiplier
γ	width of Gaussian kernel function
σ_j	standard deviation of the jth variable/feature

such as nuclear, aircraft engines or production-related processes, such as those that exist within chemical process plants, relatively little exists for application to chillers or other vapor compression equipment, especially those from the viewpoint of important sensors.

Commonly recognized categorization of FDD methodology is that based on model, quantitatively or qualitatively, and that based on process history (or data driven), as Katipamula and Brambley (2005a,b) stated. Bendapudi and Braun (2002) provide a detailed list of available quantitative models, especially dynamic models for vapor compression equipment. Qualitative physics-based and rule-based systems belong to the qualitative model-based category, such as expert systems (Kaler, 1990; Grimmelius et al., 1995; Kaldorf and Gruber, 2002), rules derived from first principles (Gordon et al., 1995; Brambley et al., 1998), bond graphs (Ghiaus, 1999) and case-based reasoning (Dexter and Pakanen, 2001), etc. The data driven methodology (Tassou and Grace, 2005; Liang and Du, 2007; Han et al., 2010) is based solely/mainly on process history and contains models of black box or gray box, where pattern recognition techniques are often employed to develop the relationship between inputs and outputs, and into which the machine learning methodology to be presented in this study falls, thanks to the data-rich nature of chillers and the dedicated experiments by Comstock and Braun (1999a,b).

Feature selection (FS) is frequently used as a preprocessing step to machine learning. It is a process of choosing a subset of original features so that the feature space is optimally reduced according to a certain evaluation criterion. FS has been a fertile field of research and development since 1970's and proven to be effective in removing irrelevant and redundant features, increasing efficiency in learning tasks, improving learning performance like predictive accuracy, and enhancing comprehensibility of learned results (Blum and Langley, 1997; Dash and Liu, 1997; Kohavi and John, 1997). Making use of FS technique, together with machine learning knowledge, this study aims to search preliminarily for relatively important sensors/features for chiller AFDD.

As a novel machine learning method, support vector machine (SVM) (Cristianini and Taylor, 2000) is a powerful tool for solving the practical problems always characterized by nonlinearity, high dimension, local minima and/or small sample. It was first suggested by Vapnik in the 1960s and has recently become an area of intense interests and research. Besides the advantages just mentioned, the main reasons for considering SVM as an AFDD tool, are: 1) the purpose of this study is concentrated on the important sensors or features in chiller AFDD, not AFDD itself; 2) SVM proved to be the most suitable technique for chiller application in Choi et al. (2005) and also in our study (Han et al.,

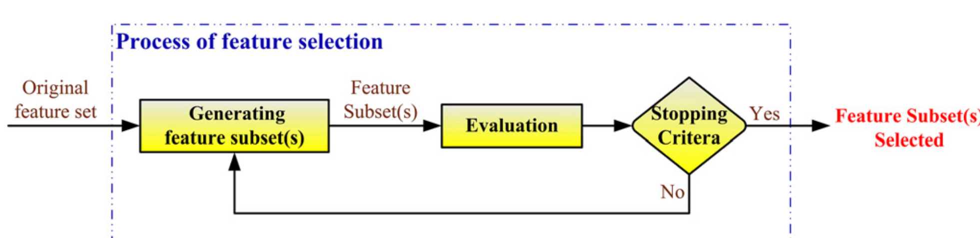

Fig. 1 – Illustration of Feature Selection (FS).