Reliability Engineering and System Safety 115 (2013) 124–135

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

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

Failure diagnosis using deep belief learning based health state classification

Prasanna Tamilselvan, Pingfeng Wang*

Industrial and Manufacturing Engineering Department, Wichita State University, Wichita, KS 67208, USA

ARTICLE INFO

Article history:
Received 22 March 2012
Received in revised form
19 February 2013
Accepted 21 February 2013
Available online 14 March 2013

Keywords:
Fault diagnosis
Artificial intelligence in diagnosis
Classification
Deep belief networks

ABSTRACT

Effective health diagnosis provides multifarious benefits such as improved safety, improved reliability and reduced costs for operation and maintenance of complex engineered systems. This paper presents a novel multi-sensor health diagnosis method using deep belief network (DBN). DBN has recently become a popular approach in machine learning for its promised advantages such as fast inference and the ability to encode richer and higher order network structures. The DBN employs a hierarchical structure with multiple stacked restricted Boltzmann machines and works through a layer by layer successive learning process. The proposed multi-sensor health diagnosis methodology using DBN based state classification can be structured in three consecutive stages: first, defining health states and preprocessing sensory data for DBN training and testing; second, developing DBN based classification models for diagnosis of predefined health states; third, validating DBN classification models with testing sensory dataset. Health diagnosis using DBN based health state classification technique is compared with four existing diagnosis techniques. Benchmark classification problems and two engineering health diagnosis applications: aircraft engine health diagnosis and electric power transformer health diagnosis are employed to demonstrate the efficacy of the proposed approach.

1. Introduction

Effective health diagnosis provides multifarious benefits such as improved safety, improved reliability and reduced costs for operation and maintenance of complex engineered systems. Research on real-time diagnosis and prognosis which interprets data acquired by smart sensors and distributed sensor networks, and utilizes these data streams in making critical decisions advances significantly across a wide range of applications [1–3]. Maintenance and life-cycle management activities, which constitute a large portion of overhead costs in many industries, would greatly benefit from these advances [4–9]. In manufacturing and service sectors, unexpected breakdowns can be prohibitively expensive since they immediately result in lost production, failed shipping schedule, and poor customer satisfaction. In order to reduce and possibly eliminate such problems, it is necessary to accurately assess current stated of system degradation through effective health diagnosis. Two major research areas have tried to address these challenges: reliability and condition monitoring. Although reliability and condition monitoring are seemingly related, reliability focuses on population-wide characteristics while condition monitoring deals with component-specific properties. Reliability analysis is generally performed based on time-to-failure data [10–13] or physics-based models (e.g., fatigue, wear, and corrosion) [14,15]. In contrast, condition monitoring research uses sensory information from functioning systems to assess their degradation states. Continuous condition monitoring and real-time failure diagnosis using sensory data acquired from smart sensors not only notify performance degradation of system components at both early and advanced stages of damages, but enable failure prognosis to facilitate crucial decision-makings on system reliability and safety improvements [16–18]. A wide range of practical applications for condition monitoring and failure diagnosis have been reported in the literature and some of the successful ones include condition monitoring of bearings [19–21], machine tools [22], transformers [23], engines [24], and turbines [25]. Despite the success, effective diagnosis of current health state based on heterogeneous sensory data from multiple sensors is still an intricate problem and remains as a major challenge for condition monitoring techniques to be applied on complex engineered systems, mainly due to high system complexity and sensory data heterogeneity. Thus, one of the most important tasks in multi-sensor health diagnosis is to develop diagnostic approaches which can effectively handle multidimensional heterogeneous sensory signals and accurately classify different health states based on these acquired sensory signals.

Despite the challenges in system health diagnosis, there is another isolated group of research specific to pattern classification in image processing, which mainly specializes on classification processes. System health diagnosis with heterogeneous

Abbreviations: BNN, back-propagation neural network; DBN, deep belief networks; HS, health state; MD, Mahalanobis distance; PHM, prognostics and health management; RBM, restricted Boltzmann machine; SVM, support vector machine; SOM, self-organizing map

* Corresponding author. Tel.: +1 316 978 5910, fax: +1 316 978 3742.
E-mail addresses: pxtamilselvan@wichita.edu (P. Tamilselvan), pingfeng.wang@wichita.edu (P. Wang).

0951-8320/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2013.02.022

Nomenclature

$\mathbf{x}_i$	p-dimensional vector
c_i	ith class label
$\boldsymbol{\mu}_j$	mean vector of the training data
$\mathbf{S}_j$	variance matrix of the training data
w_{ij}	synaptic weight between the ith and the jth neurons
s_i	state of the ith neuron
$P(\cdot)$	probability distribution function
b_i	bias of the ith neuron
h_i	state of the ith neuron in hidden layer
v_i	state of the ith neuron in visible layer
m	momentum
n	epoch number

multidimensional sensory data is similar with the pattern classification problem with a high dimensionality of image data. In both cases, machine learning techniques have been dominant approaches and the learning complexity grows exponentially with the increase in heterogeneity and dimensionality of acquired sensory data [26]. In the past decade, pattern classification techniques have moved into a new platform of learning procedure called deep machine learning [26]. Analysis of the similarity between health diagnosis and pattern classification in these different applications motivates the emergence of a perfect collaboration in their learning techniques. There lies a great potential and also a critical need to utilize the advantages of deep machine learning techniques to address the challenges faced in system health diagnosis. However, advantages of evolving deep machine learning process have not been employed in current condition monitoring and health diagnosis research. Thus, this study proposes an efficient way to utilize the benefit of deep machine learning process to handle the complexity of sensory signals for the application of structural health diagnosis.

This paper presents a novel multi-sensor health diagnosis method using deep belief network (DBN). DBN has recently become a popular approach in machine learning for its promised advantages such as fast inference and the ability to encode richer and higher order network structures. The DBN works based on the restricted Boltzmann machine (RBM) and learns layer by layer of a deep network structure. The proposed diagnosis methodology can be structured in three consecutive stages: first, defining health states and preprocessing sensory data for DBN training and testing; second, developing DBN based classification models for the diagnosis of predefined health states; third, validating DBN classification models with testing sensory dataset. Health diagnosis using DBN based health state classification technique is compared with four existing diagnosis techniques: SVM, BNN, SOM, and MD classifier. Benchmark classification problems and two engineering health diagnosis applications: aircraft engine health diagnosis and electric power transformer health diagnosis are employed to demonstrate the efficacy of the proposed approach. The rest of the paper is organized as follows: Section 2 presents the related work of health diagnosis with existing state of the art classification techniques; Section 3 details the proposed health diagnosis approach with DBN based classification; Section 4 demonstrates the developed diagnosis approach with case studies; and Section 5 summarizes the presented research and the future work.

2. Related work

Due to the complexity of system degradation characteristics and potential heterogeneity of sensory signals, multi-sensor health diagnosis of complex engineered systems remains as an intricate problem. Consequently, machine learning techniques and statistical inference techniques are often employed to solve this problem. Significant advances have been achieved in applying classification techniques based on machine learning [27–34] or statistical inferences [35–37], which resulted in a number of state-of-the-art health state classification methods, such as back-propagation neural network (BNN) [27–30], self-organizing maps (SOM) [31], support vector machine (SVM) [28,32–34], and Mahalanobis distance (MD) [32,35]. This section surveys the current literature and briefly discusses the working principle and capabilities of the different existing classification techniques.

In general, machine learning based diagnosis techniques can be broadly classified into supervised and unsupervised learning approaches. This paper focuses on the supervised learning process in diagnostic classification. The supervised learning is the process of learning a relationship between input values and desired target labels in the form of set of patterns. The error values are evaluated and given as feedback to the learning model to get potential optimal solutions. The learnt relationship/function from training dataset is used as a classifier model to predict unlearnt dataset with unknown patterns. Unlike supervised learning process, unsupervised learning process does not utilize labeled training data. The BNN and the SOM are two representative artificial neural network type diagnosis techniques that are based on supervised learning and unsupervised learning respectively. The BNN possesses a basic neural network structure with three different types of layers: the input layer, the output layer and the hidden layers [27], and is generally trained through optimizing synaptic weights and biases of all neurons till a desired classification rate is obtained. Using BNN, the health diagnosis problem is solved as a health state prediction problem using trained neural networks. Different with the BNN, the SOM is an unsupervised learning technique working based on neurons that determines a closest best-matching unit distance to the input vector [31], and uses it to construct class boundaries graphically on a two dimensional map. The BNN and the SOM have been used as state classification methods for different health diagnosis applications [27]. The main drawback of the BNN and SOM is the over-fitting of the data to the classification model and leading to significant error values in complex scenarios.

The SVM is a state-of-the-art technique for multi-dimensional classification based on supervised learning. The SVM organizes input data D into two sets of vectors in a p-dimensional space as

$$D=\{(\mathrm{x}_i,c_i)|\mathrm{x}_i\in R^p,\quad c_i\in\{0,1,2\},\ \ i=1,2,\ldots,p\},\tag{1}$$

where c_i is the ith class label (e.g., 0, 1, or 2) indicating the class to which data point $\mathbf{x}_i$ belongs. Each $\mathbf{x}_i$ is a p-dimensional real vector, shows the preprocessed p-dimensional sensory data. With the organized input data, the SVM constructs hyper-planes with maximum margins to divide data points with different c_i values [33]. A hyper-plane can be written as a set of points $\mathbf{x}$ satisfying

$$\mathbf{w}\cdot\mathbf{x}-b=0,\tag{2}$$

where vector $\mathbf{w}$ is a normal vector that is perpendicular to the hyper-plane. The parameter $b/\|\mathbf{w}\|$ determines the offset of the hyper-plane from the origin along the normal vector $\mathbf{w}$. We want to choose the $\mathbf{w}$ and b to maximize the margin, or distance between the parallel hyper-planes of the margin. The optimization