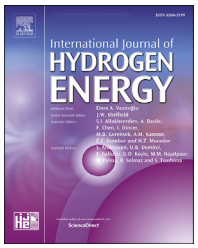
Available online at www.sciencedirect.com

ScienceDirect

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



A discrete hidden Markov model fault diagnosis strategy based on K-means clustering dedicated to PEM fuel cell systems of tramways

Jiawei Liu [a], *Qi Li* [a,*], *Weirong Chen* [a], *Taiqiang Cao* [b]

[a] School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, Sichuan Province, China
[b] School of Engineering and Electric Information, Xihua University, Chengdu 610039, Sichuan Province, China

ARTICLE INFO

Article history:
Received 1 October 2017
Received in revised form
11 April 2018
Accepted 22 April 2018
Available online 30 May 2018

Keywords:
Fuel cell tramways
Fault diagnosis
Discrete hidden Markov model
K-means clustering
Scalar quantization

ABSTRACT

To solve the fault classification problems of fuel cell (FC) various health states for tramways, a discrete hidden Markov model (DHMM) fault diagnosis strategy based on K-means clustering is proposed. In this paper, the K-means clustering algorithm is used to filter the sample points which aren't consistent with the actual class labels. The Lloyd algorithm is employed to quantify the sample vector sets and obtain the discrete code combination of training samples and test samples. The Baum-Welch algorithm and forward-backward algorithm are respectively presented to train and deduce the DHMM. The classification results show that the six concerned faults can be detected and isolated. The targeted fault types include low air pressure, deionized glycol high inlet temperature, deionized humidification pump low pressure, deionized glycol outlet temperature signal voltage overrange, normal state and hydrogen leakage. The fault recognition rates with the novel approach are at best 94.17%.

Introduction

Fuel cell converts the chemical energy of hydrogen and oxygen into electrical energy by electrochemical device [1–3]. It has the advantages of high energy conversion rate, low noise, fast dynamic response and modular design. It has broad prospects in distributed power generation, city tramways [4,5], city buses, electric vehicles and other transportation areas [6–8]. However, the fuel cell stack system (FCSS) is prone to malfunction in practical operation. The slight faults may cause reduction of output power and shutdown of stack. The serious failures can permanently damage the FC [9–11]. Therefore, fault detection and isolation (FDI) of fuel cell system has become one of the most important issues [12–14].

In recent years, artificial intelligence (AI) technologies have been widely used in the fault diagnosis of proton exchange membrane fuel cell (PEMFC) systems [15]. In Refs. [16,17], a cascade controller of the extended state observer (ESO) has been proposed for regulating the oxygen excess ratio of the PEMFC air-feed system to its desired value, based on the second-order sliding mode technique. The problem of fault diagnosis has been firstly considered in Ref. [18] for a class of nonlinear systems and successfully applied to a fuel cell air feed system. Lei Mao et al. [19,20] have proposed a novel sensor selection algorithm and its performance in PEMFC online diagnosis is investigated. The results show that the

* *Corresponding author.*
E-mail address: liqi0800@163.com (Q. Li).
https://doi.org/10.1016/j.ijhydene.2018.04.163

proposed algorithm can provide more efficient sensor selection results with less computational time. In Refs. [21,22], a general framework of fault diagnosis has been proposed to deal with multiple sensor measurements. Results show that features selected using singular value decomposition (SVD) can provide more effective classification results. E. Pahon et al. [23] have proposed a novel signal-based approach for fault diagnosis of PEMFC, which focuses on the detection of a high air stoichiometry (HAS) fault. The wavelet transform (WT) is used to diagnose quickly an oversupply of air to the PEMFC system. In order to describe the nonlinear voltage and current characteristics PEMFCs, D. Ritzberger et al. [24] have proposed a fault diagnosis method based on polynomial nonlinear autoregressive models with exogenous inputs (NARX) and the Volterra series. Ali Mohammadi et al. [25] have presented a 2 steps diagnosis approach for PEMFC. First, a 3D fault-sensitive model is built to simulate different faults (drying out and flooding) with 2 severity levels ("flooding", "drying", "high flooding" and "high drying"). Then, a two-layers feed-forward artificial neural network (ANN) has been developed to localize each fault in 9 considered segments within a single cell. The reservoir computing (RC) has been introduced into the field of FC diagnosis [26]. The targeted fault types include CO poisoning, low air flow rate, cooling circuit failure and system aging. Xingwang Zhao et al. [27] have proposed a fault diagnosis method based on multi-sensor signals and principal component analysis (PCA). The serious system fault and a single sensor failure have been successfully diagnosed and distinguished. Support vector machine (SVM) and designed diagnosis rules have been used to achieve FDI in real-time [28]. Four concerned failures have been diagnosed, including membrane drying, low air stoichiometry, low pressure fault and high pressure fault. The wavelet transform modulus maxima (WTMM) and pattern recognition methods have been combined to realize the fault diagnosis of PEMFC [29,30]. The continuous wavelet transform (CWT) and multifractal have been employed to calculate the singularity spectrum of the voltage signal. The SVM and K-nearest neighbor (KNN) algorithm have been used to classify the spectrum. The results show that the method can distinguish the faults such as cathode flow rate, gas pressure, cooling circuit and CO poisoning. The PEMFC electrochemical characteristic information has been extracted from the output terminal voltage signal (OTVS) by discrete wavelet decomposition [31]. Mona Ibrahima et al. [32] have presented discrete wavelet transform (DWT) to identify the membrane drying and flooding faults. The research shows that the DWT is more effective than the CWT in the field of fault diagnosis for fuel cell. Cédric Damour et al. [33] have proposed a novel empirical mode decomposition (EMD) based diagnosis approach dedicated to PEMFCs. The flooding and drying faults can be diagnosed by comparing the voltage contribution between the first and ninth intrinsic mode functions (IMFs). In order to find an online diagnostic tool for PEMFC, literature [34,35] have compared various types of PEMFC diagnostic methods in terms of diagnostic accuracy and computational complexity. PCA, Fisher discriminant analysis (FDA), kernel principal component analysis (KPCA) and kernel Fisher discriminant analysis (KFDA) have been used to extract features from the PEMFC voltage. Gaussian mixture model (GMM), KNN and SVM have been employed to make decisions. The experimental results show that the strategy based on FDA and SVM has higher performance and less computational cost than the other approaches.

However, most of the research work is focused on the small and medium size applications. The research on the fault diagnosis of the high power fuel cell system for the tramways is almost in the blank state. Electrochemical impedance spectroscopy (EIS) requires a high degree of professional knowledge with time-consuming and poor visibility. The kernel function of SVM is limited by the Mercer condition. The training time of relevant vector machine (RVM) increases rapidly with the increase of training samples. The sample size of neural network (NN) is large and easy to fall into "local minimum" and "over fitting". The fault samples can only be classified into several subclasses by clustering. The expert system has poor ability to acquire empirical knowledge and can't meet the requirements of real-time performance. The significant information may be lost with PCA.

Hidden Markov model (HMM) is a statistical modeling method for nonstationary time series [36,37]. It has the advantages of small training samples and interpretable feature. It can simultaneously reflect the randomness and latent structure of objects. It is applicable to modeling problems of complex dynamic nonlinear systems. HMM has been successfully used in many fields, such as network status analysis, state detection of electrical equipment, speech recognition and computer vision [38–41]. However, as far as we know, little literature studies the practical application of the Markov theory to identify the PEMFC faults. Bayesian-score (K2) and Markov chain Monte Carlo (MCMC) algorithms have been combined to construct the graphical-probability structure for fault diagnosis of FC [42]. Four types of faults in PEMFC are considered.

In this paper, a DHMM fault diagnosis strategy based on K-means clustering (DHMM-K) dedicated to PEM fuel cell systems of tramways is proposed for the first time. It has higher recognition accuracy and better expansibility for a variety of health states. The K-means clustering algorithm is used to eliminate singular sample points. The Lloyd method is employed to quantify the sample vector sets and obtain the discrete code combination of training samples and test samples. The Baum-Welch algorithm and the forward-back algorithm are respectively adopted to train and infer the DHMM. The experimental data from the tramway are proposed to verify the feasibility of the novel strategy. The advantages of the proposed approach are further proved by comparison with one-against-one (OAO) SVM.

Analysis of typical PEMFC faults for tramways

Fuel cell system structure in the tramway

FC/Super capacitor hybrid tramway is developed by Southwest Jiaotong University and CNR Tangshan Railway Vehicle Co., Ltd., as shown in Fig. 1. The configurations for the tramway have been listed in Table 1.

The detailed structure of PEMFC system has been shown in Fig. 2, which includes PEMFC subsystem, air supply subsystem and cooling cycle subsystem. The 150 kW Fcvelocity-HD6