Journal of Power Sources 378 (2018) 646–659

Contents lists available at ScienceDirect

Journal of Power Sources

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

Data-driven simultaneous fault diagnosis for solid oxide fuel cell system using multi-label pattern identification

Shuanghong Li[a,b], Hongliang Cao[c,d,*], Yupu Yang[a,b]

[a] Department of Automation, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, PR China
[b] Key Laboratory of System Control and Information Processing, Ministry of Education, Shanghai 200240, PR China
[c] College of Engineering, Huazhong Agricultural University, No. 1, Shizishan Street, Hongshan District, Wuhan 430070, PR China
[d] Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River, Ministry of Agriculture, Wuhan 430070, PR China

HIGHLIGHTS

- A data-driven fault diagnosis approach is designed for SOFC systems.
- The approach achieves high diagnosis accuracy for unseen simultaneous faults.
- The framework only requires single fault data for the ML-SVM training.
- Simultaneous fault symptoms can be deducted by the FITA.

ARTICLE INFO

Keywords:
SOFC system
Data-driven
Multi-label
Pattern identification
Simultaneous faults

ABSTRACT

Fault diagnosis is a key process for the reliability and safety of solid oxide fuel cell (SOFC) systems. However, it is difficult to rapidly and accurately identify faults for complicated SOFC systems, especially when simultaneous faults appear. In this research, a data-driven Multi-Label (ML) pattern identification approach is proposed to address the simultaneous fault diagnosis of SOFC systems. The framework of the simultaneous-fault diagnosis primarily includes two components: feature extraction and ML-SVM classifier. The simultaneous-fault diagnosis approach can be trained to diagnose simultaneous SOFC faults, such as fuel leakage, air leakage in different positions in the SOFC system, by just using simple training data sets consisting only single fault and not demanding simultaneous faults data. The experimental result shows the proposed framework can diagnose the simultaneous SOFC system faults with high accuracy requiring small number training data and low computational burden. In addition, Fault Inference Tree Analysis (FITA) is employed to identify the correlations among possible faults and their corresponding symptoms at the system component level.

1. Introduction

Solid oxide fuel cells (SOFC) system directly generates electric power from hydrogen. The generated electric power has a high operation temperature, zero greenhouse gas emission, and hydrogen impurity tolerance [1–5]. These properties enable SOFC systems to locate downtown areas and densely populated regions for residual, commercial, and electrical usages. However, an SOFC system may take place on various kinds of faults resulting to system accidents and economic losses. In SOFC real operations, faults may occur on any parts at any time, such as fuel leakage fault, air compressor fault or sensors failures etc. These malfunctions may lead to performance losses, irreversible degradation and even system failure. In order to improve the stability and durability for the SOFC system, plenty of fault diagnosis methods have been proposed [6,7].

Since fault prediction is regarded as one of the best ways for preventing SOFC system accidents, many SOFC system fault detection and diagnosis methods have been studied recently [8–28]. The method of these works can be sorted into two main classes: model-based fault diagnosis methods and data-based fault diagnosis approaches. Model-based fault diagnosis methods, wherein SOFC models, also called physical models, are built; the instantaneous remoteness between the actual SOFC behavior and the expected healthy behavior are then computed; and finally, through the residue analysis, faults can be detected [8–13]. Among them, Greco et al. [8] analyzed the effects of fuel processor faults in an SOFC system by using the whole system model developed by first principles. Sun et al. [10] presented a hierarchical structure for fault detection and fault-tolerant control by using a simulated model of an SOFC plant. Based on an improved equivalent

* Corresponding author. College of Engineering, Huazhong Agricultural University, No. 1, Shizishan Street, Hongshan District, Wuhan 430070, PR China.
E-mail address: hongliangcao@mail.hzau.edu.cn (H. Cao).

https://doi.org/10.1016/j.jpowsour.2018.01.015
Received 28 October 2017; Received in revised form 20 December 2017; Accepted 7 January 2018

Nomenclature

Acronyms

SOFC	Solid Oxide Fuel Cell
PCA	Principal Component Analysis
FITA	Fault Inference Tree Analysis
SL	Single-Label
ML	Multi-Label

Roman symbols

C	Specific heat capacity, kJ kg^{-1} K^{-1}
E	Testing data matrix
F	Faraday constant = 96485 C mol^{-1}
h_{gs}	Convection heat transfer coefficient, kW m^{-2} K^{-1}
i	Current density, A cm^{-2}
N	Control volume mole number, mol
Q	Heat transfer, W
R_i	Reaction rate of species i, mol s^{-1}
R	Train data matrix
S	Surface area of the heat transfer, m^2
T	Absolute temperature, K
W	Work, W
X	Data matrix
X_i	Species mole fractions i

Greek symbols

α	Transfer coefficient
ρ	Density, kg m^{-3}
ε	Fuel fault sensitive factor
δ	With of RBF

Subscripts and superscripts

act	Activation
con	Concentration
cond	Conduction
conv	Convection
f	Fault
H_2	Hydrogen
P	Pressure
S	Solid
OCV	Irreversible open circuit voltage
O_2	Oxygen
Rec	Recall
L	Limit

circuit model, Huang et al. [11] conducted an impedance diagnosis of metal-supported SOFCs. By using the model-based fault diagnosis method, a new control strategy that tolerates the SOFC system faults is proposed in Ref. [13] to diagnose the fault types.

Since the processes become more and more complicated, the physical principles of the systems are difficult to obtain in practice. The model-based fault diagnosis methods are therefore hard to obtain. Compared with the well-established model-based methods, which demands qualitative knowledge concerning the systems, the methods driven by data [25], taking advantage of the information from the substantial process merely rely on the historical recorded data sets, are receiving more and more attentions. The production process monitoring and fault diagnosis schemes with data-based fault identifying techniques have prospered and begun to be applied in the industry to protect the production system from disastrous accidents and keep stability [29–34]. Data-based fault diagnosis approaches for SOFC system involve fault-tree algorithms [14,15], Principal Component Analysis (PCA) methods [16], Artificial Neural Networks (ANN) [17], and Support Vector Machine (SVM) [18,19]. Polverino et al. [14] designed an on-field diagnostic algorithm for SOFC systems. The diagnosis design phase relies on the fault tree analysis, which identifies the correlations among possible faults and their corresponding symptoms at the system components level. Steiner et al. [15] presented an overview of the use and the contribution of fault tree analysis to SOFC and PEFC diagnoses. Murshed et al. [16] designed a monitoring system for SOFC systems by using a hybrid PCA method. The development of recurrent neural networks and neural network classifier models for modeling and diagnosing SOFC stacks was reported in Ref. [17]. Costamagna et al. [18] proposed the use of a quantitative model for SOFC systems with SVM to detect and classify possible faults. An overview of all the proposed SOFC diagnosis methods was presented in Ref. [6].

All the studies have made certain contributions to detect and classify single or multi-faults. They have tried to solve two kinds of fault diagnosis problems. The first one is the fault detection which involves the use of binary diagnosis to identify the existence of faults [16]. SOFC has a number of independent faults, which are also referred to as single faults. As a result, the second kind of works, also called as multiple fault diagnosis, tries to solve multiclass diagnosis problems. This diagnostic tool returns the most probable single fault based on model-system residuals (model-based method) [8–14] or observed data (data-driven method) [15–19]. However, several single faults, also called the simultaneous faults, may also appear at the same time. To the best of our knowledge, no studies have yet to try to diagnose the SOFC system simultaneous faults. Simultaneous-fault diagnosis is a method based on the observed patterns or data for the detection of single faults that occur simultaneously.

Furthermore, simultaneous fault diagnosis for industrial systems have faced two challenges: 1) the features of the multiple faults are mixed or combined into one pattern which makes accurate diagnosis difficult; 2) the acquisition of a large sample data set of simultaneous faults is expensive due to the high number of combinations of single faults, resulting in many possible classes of simultaneous-fault training patterns. Recent literatures have showed that these simultaneous fault diagnosis problems were solved using data-driven Multi-Label (ML) classification [30–34]. The ML technique, which uses SVM as the learning algorithm to arrange a simultaneous-fault diagnosis system, was first used in Refs. [30] and [31]. Pooyan et al. [32] designed a novel ML–SVM approach based on multiple regulation parameters for simultaneous-fault classification in a dew point process. Through an illustrative application of automotive engine diagnosis, Vong et al. [34] proposed a new simultaneous-fault diagnosis framework by using pairwise probabilistic ML classification. Moreover, these studies, have introduced two label approaches on facing a classification or fault diagnosis problem: the Single-Label (SL) approach, which classifies a set of patterns into a univocal class, and the ML approach, assigns each input data to more than just one class. The ML-based pattern identification is currently considered as the best solution of simultaneous-fault diagnosis to overcome the drawbacks of artificially defined classes. In addition, ML-based pattern identification is also capable of diagnosing medium or even large data processing problems. Another advantage of ML approach over other methods is that the ML approach does not require expensive simultaneous fault patterns on training data sets and it can also train diagnosis systems by only using single classes through the decomposition of simultaneous class patterns into corresponding combinations of single fault classes [30,31]. By this way, the occurrence of a trained fault can be detected in the future even if the fault comes together with other classes of faults.

This research is inspired by the above-mentioned studies that have focused on the use of ML pattern identification technology to diagnose the simultaneous faults in simulated SOFC stand-alone system. The simultaneous-fault diagnosis framework includes two components such as feature