2014 IEEE/ASME International Conference on
Advanced Intelligent Mechatronics (AIM)
Besançon, France, July 8-11, 2014

Fault diagnosis and novel fault type detection for PEMFC system based on Spherical-Shaped Multiple-class Support Vector Machine

Zhongliang Li, Stefan Giurgea, Rachid Outbib, and Daniel Hissel

Abstract— **In this paper, a data-based strategy is proposed for PEMFC (polymer electrolyte membrane fuel cell) diagnosis. In the strategy, the feature extraction method *Fisher Discriminant Analysis* (FDA) is used firstly to extract the features from individual cell voltages. After that, the classification method *Spherical-Shaped Multiple-class Support Vector Machine* (SSM-SVM) is used to classify the extracted features to various classes related to health states. The potential novel failure mode can be detected in the procedure. Experiments on a 40-cell stack are dedicated to verify the approach.**

I. Introduction

Increasing environment and resource issues motivate the development and commercialization of fuel cell technologies. PEMFC (polymer electrolyte membrane fuel cell) is one of the most promising fuel cells, especially for mobile applications. To improve the reliability and durability of PEMFC, much effort is being made to the research of the degradation mechanisms of materials, and the design and assembly of fuel cells. Apart from the researches of internal improvements, the topic of fault diagnosis for the PEMFC system is currently receiving considerably increasing attention. It is considered that an efficient fault diagnosis strategy can help PEMFC (stack) operating in a relatively optimal condition, and thus mitigate the performance degradation of fuel cells.

Due to the roadblock of identification of fuel cell inner parameters, a sufficiently accurate and diagnosis oriented model is usually difficult to obtain [1]. Data-based fault diagnosis methodologies have been drawing more and more attentions. Some data-based methods have been proposed for PEMFC diagnosis during the recent years [2] [3] [4]. Although some preliminary results have been obtained by these attempts, more effort is still deserved in this direction. For instance, from the practical point of view, the performance of the studied diagnostic methods, such as diagnosis accuracy, computation cost, should be evaluated seriously. Moreover, most of the diagnostic tools in the literature concern some specific fault categories. The performance of multi-fault isolation need to be investigated and verified.

Z. Li and R. Outbib are with the LSIS Laboratory, UMR CNRS 6168, University of Aix-Marseille, 13397 Marseille Cedex 20, France (e-mail: zhongliang.li@lsis.org; rachid.outbib@lsis.org). Z.li is also with FEMTO-ST/ENERGY (UMR CNRS 6174) and FCLAB (CNRS 3539), Belfort 90010 Cedex, France

S. Giurgea is with the University of Technology Belfort-Montbéliard (UTBM) and FCLAB (CNRS 3539), Belfort 90010 Cedex, France (e-mail: stefan.giurgea@utbm.fr).

D. Hissel is with the University of Franche-Comté (UFC), with FEMTO-ST/ENERGY (UMR CNRS 6174) and FCLAB (CNRS 3539), Belfort 90010 Cedex, France (e-mail: daniel.hissel@univ-fcomte.fr).

In the domain of data-based diagnosis, pattern classification techniques have been widely used for data-based diagnosis. The classification based diagnostic procedure usually proceeds in two steps. Firstly, an empirical classifier is established from prior knowledge and history data, which is considered as a training process. Secondly, the real-time data are processed by the classifier in order to determine whether and which faults occur. In our previous work [5], classification techniques, *Fisher Discriminant Analysis* (FDA) and *Support Vector Machine* (SVM), have been adopted to detect the faults, named flooding and membrane drying. The approach is highlighted by its high diagnosis accuracy and low computation cost, and thus can be seen as a candidate of online diagnostic tool of PEMFC system.

Nevertheless, the main disadvantage of the conventional classification methods is: the classifiers trained using the pre-existing data can only be used recognize the known failure modes. An arbitrary example would be classified to a known class even if it belongs to a new cluster which strongly differs from the training dataset. For the industrial systems, such as fuel cell system, it is usually not possible to get the data in all the failure modes at the training stage. It can be imagined that the data of unseen failure modes would be always misclassified. The capability of the detection of potential novel fault type, seems to be needful in such cases.

In this paper, we propose a data-based fault diagnosis strategy for PEMFC system. Similarly as the approach of our previous work [5], the individual cell voltages are employed as the original variables for diagnosis, and FDA is used to extract the features for classification. A new classifier, named Spherical-Shaped Multiple-class Support Vector Machine (SSM-SVM) is adopted to classify the features to various classes (fault free state and various fault states). At the same time, the data from potential novel failure mode could be detected.

The contributions of this paper are summarized as follows:

- Multiple faults, which involve various parts of PEMFC system, can be detected and isolated with a high accuracy.
- The diagnostic approach can detect the data from a potential novel cluster without supplemental procedures.
- The computation cost can certify the real-time implementation of the approach.

The rest of the paper is organized as follows: In section 2, the diagnosis strategy is expounded. The whole framework is given firstly. Following that, FDA, SSM-SVM and the diagnosis decision rules are explained. In section 3, the strategy is tested using the experimental data of a 40-cell

PEMFC stack. Finally, the conclusion and future work are summarized in section 4.

II. Diagnostic strategy

A. Framework of the strategy

Basically, the proposed diagnostic strategy includes two stages: off-line training and on-line performing. Fig. 1 outlines the framework of the approach.

In the off-line training stage, the models of FDA and SSM-SVM are trained successively based on the training database (constructed by historical samples of cell voltages) that distributes in various classes. In the on-line performing stage, the real-time sample (cell voltages) is firstly processed by the trained FDA model. Through the procedure, features can be extracted from raw data. After that, with the aid of trained SSM-SVM model, the features are classified either to a certain known class to get the diagnostic decision, or to a potential novel cluster. The data which are justified to the potential novel cluster will be supplied to the training database for retraining the diagnostic models.

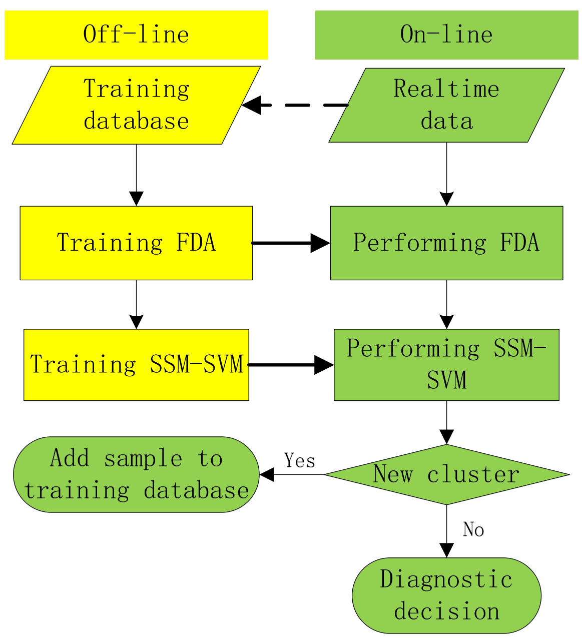

Fig. 1. Flowchart of the proposed diagnostic strategy

B. Problem description

The diagnostic problem can be abstracted mathematically as follows: suppose that we have a training dataset of N H-dimensional samples $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$, which are distributed in C classes denoted by $\omega_1, \omega_2, \ldots, \omega_C$. The sample number in ω_i is N_i. FDA and SSM-SVM models are trained based on the training dataset. Through the trained models, a real-time sample $\boldsymbol{x}$ can be classified to a defined class ω_i, $i = 1, \ldots, C$ or a novel cluster denoted by ω_{novel}.

C. FDA

FDA is a technique developed for feature extraction or dimension reduction in hope of obtaining a more manageable classification problem. Through FDA, the original H-dimensional samples are projected to a L-dimensional space $(L < H)$ with L unit projecting vectors (H-dimension):

$$\boldsymbol{z}_n = [\boldsymbol{w}_1^T \boldsymbol{x}_n, \boldsymbol{w}_2^T \boldsymbol{x}_n, \ldots, \boldsymbol{w}_L^T \boldsymbol{x}_n]^T \quad n = 1, \ldots, N \tag{1}$$

where $\boldsymbol{z}_n$ is the *projected vector* corresponding to $\boldsymbol{x}_n$, and $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, $\ldots$, $\boldsymbol{w}_L$ are *projecting vectors*. The *projected vectors* are constructed by the *features* that locate in the *feature space*.

The objective of training FDA is to find the *projecting vectors*, which make the *projected vectors* in the same class are concentrated, while those in different classes are separated as possible [6].

Avoiding the detailed theoretical deducing process (which can be found in [6] and [7]), the seeking of the *projecting vectors* is converted to the following eigen vectors problem

$$\boldsymbol{S}_b \boldsymbol{w}_i = \lambda_i \boldsymbol{S}_w \boldsymbol{w}_i \; i = 1, \ldots, L \text{ and } \lambda_1 \geq \ldots, \geq \lambda_L > 0 \tag{2}$$

where $\boldsymbol{S}_w$ and $\boldsymbol{S}_b$ are *within-class* covariance matrix and *between-class* covariance matrix, which are given

$$\boldsymbol{S}_w = \sum_{i=1}^{C} \sum_{\boldsymbol{x}_n \in \omega_i} (\boldsymbol{x}_n - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_n - \bar{\boldsymbol{x}}_i)^T$$

and

$$\boldsymbol{S}_b = \sum_{i=1}^{C} N_i (\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})(\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})^T$$

where $\bar{\boldsymbol{x}}_i$ is mean vector in class ω_i: $\bar{\boldsymbol{x}}_i = \sum_{\boldsymbol{x}_n \in \omega_i} \boldsymbol{x}_n / N_i$, and $\bar{\boldsymbol{x}}$ is mean vector of total dataset: $\bar{\boldsymbol{x}} = \sum_{n=1}^{N} \boldsymbol{x}_n / N$.

It is found that no more than $C-1$ of the eigenvalues are nonzero and the projecting vectors are corresponding to these nonzero eigenvalues [6]. Hence, the dimension of feature space L should satisfy the constraint

$$L \leq C - 1 \tag{3}$$

In this study, the equalization case is preferred. After training FDA, the *projected vectors* $\{\boldsymbol{z}_n\}$ corresponding to $\{\boldsymbol{x}_n\}$ are exported to the next classification step. *Projecting vectors*, $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, $\ldots$, $\boldsymbol{w}_L$, are saved for the use of performing.

D. SSM-SVM

Support vector machine (SVM) has been considered as a powerful classifier, for its superior characteristics, such as local minima can be avoided and the solutions can be represented sparsely [8]. In addition and to emphasize, compared with another widely-used classifier Artificial Neural Networks (ANN), an attractive advantage of SVM is that SVMs are less prone to over fitting problem [9]. SVM was originally designed for binary classification. By combining the basic binary classifiers, the binary SVM can be extended to multi-class SVM classification ("one-against-one", "one-against-all", "DAGSVM" [10]). Although the classifiers based on the binary SVM can classify the known classes, the capability of detecting a unseen cluster seems to be defective. As Fig. 2 (a) shows, the bounders among trained classes (Class1, Class2, and Class3) can be affirmed by a multi-class classifier based on binary SVM. According to the decision of the trained classifier, an arbitrary sample will be classified to one of the three classes. Even if the data are from a novel cluster (as Fig. 2 (b) shows).

Different from binary SVM based classification, authors of [11] proposed spherical-structured multi-class SVM. The approach finds a sphere for each class. The class-specific sphere can enclose the samples from the corresponding class while exclude those outside the class. As Fig. 2 (c) shows, the closed bounders for all the known classes can be found by training SSM-SVM. Thus, samples from a novel cluster can be detected if they are out of all the closed bounders (as Fig. 2 (d) shows).

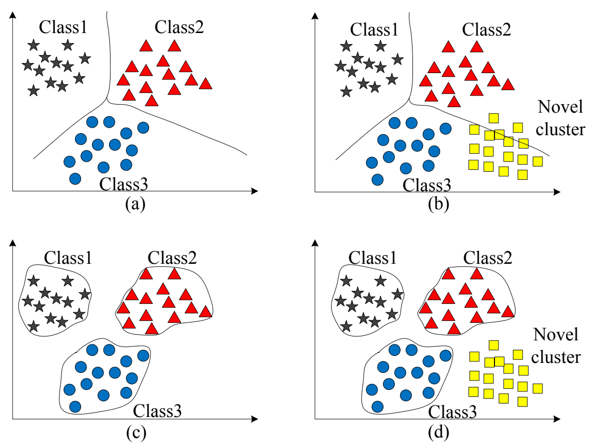

Fig. 2. Schematic diagram of conventional multi-class SVM and SSM-SVM

Following the FDA step, the training of SSM-SVM is based on the *projected vectors* $\{\boldsymbol{z}_n\}$. To solve the nonlinear classification problem, $\{\boldsymbol{z}_n\}$ are firstly projected to a high-dimensional space via a nonlinear transform Φ [11]. For ith class, the method is realized by seeking the sphere with the minimal radius in a high-dimensional space. The samples in the ith class are enclosed in ith sphere, while those outside ith class are excluded outside. This is formulated as

$$||\Phi(\boldsymbol{z}_n) - \boldsymbol{a}_i||^2 \leq R_i^2 + \xi_n^i \quad \text{if } \boldsymbol{z}_n \in \omega_i$$
$$||\Phi(\boldsymbol{z}_n) - \boldsymbol{a}_i||^2 \geq R_i^2 - \xi_n^i \quad \text{if } \boldsymbol{z}_n \notin \omega_i$$

where R_i and $\boldsymbol{a}_i$ are respectively the center and radius of the ith sphere, and ξ_n^i are the slack variables, which allow the inseparable data can be handled. The above two inequality is reduced to

$$c_n^i(||\Phi(\boldsymbol{z}_n) - \boldsymbol{a}_i||^2) - \xi_n^i \leq 0 \quad \text{and} \quad \xi_n^i \geq 0 \quad \forall n \tag{4}$$

where $c_n^i = 1$ if $\boldsymbol{z}_n \in \omega_i$ and $c_n^i \neq 1$ if $\boldsymbol{z}_n \notin \omega_i$. With the constraint conditions of (4), the following quadratic problem (QP) is formulated:

$$\min_{R_i, \boldsymbol{a}_i} R_i^2 + D \sum_n \xi_n^i \tag{5}$$

where D is a parameter controlling the penalty of misclassified points.

By introducing Lagrange multipliers $\{\alpha_n^i\}$, and the kernel function $k(\boldsymbol{z}_n, \boldsymbol{z}_m) = \Phi(\boldsymbol{z}_n)\Phi(\boldsymbol{z}_m)$[1], the following dual problem that involves $\{\alpha_n^i\}$ is obtained (see [11])

$$\min \sum_n \sum_m \alpha_n^i c_n^i \alpha_m^i c_m^i k(\boldsymbol{z}_n, \boldsymbol{z}_m) - \sum_n \alpha_n^i c_n^i k(\boldsymbol{z}_n, \boldsymbol{z}_n) \tag{6}$$

subject to

$$\sum_n \alpha_n^i c_n^i = 1 \quad \text{and} \quad 0 \leq \alpha_n^i \leq D \quad \forall i \tag{7}$$

The training process is to solve the QP problem (6). In our study, an incremental learning method, given by [12], is adapted to solve the problem efficiently.

After solving the problem, the distance from a data point $\boldsymbol{z}$ to the ith sphere center can be expressed (see [11])

$$\begin{aligned} d_i^2(\boldsymbol{z}) &= ||\Phi(\boldsymbol{z}) - \boldsymbol{a}_i||^2 \\ &= k(\boldsymbol{z}, \boldsymbol{z}) - 2\sum_n a_n^i c_n^i k(\boldsymbol{z}, \boldsymbol{z}_n) + \sum_{n,m} a_n^i c_n^i a_m^i c_m^i k(\boldsymbol{z}_n, \boldsymbol{z}_m) \end{aligned} \tag{8}$$

The radius R_i is given

$$R_i^2 = ||\Phi(\boldsymbol{z}_n) - \boldsymbol{a}_i||^2 \quad \forall \boldsymbol{z}_n \text{ with } \alpha_n^i \in (0, D) \tag{9}$$

E. Diagnostic decision rules

The criterion, which weighs the membership degree of $\boldsymbol{z}$ belonging to class ω_i is given by the authors of [11], as

$$\mu_i(\boldsymbol{z}) = \begin{cases} 0.5\left(\dfrac{1 - d_i(\boldsymbol{z})/R_i}{1 + \lambda_1 d_i(\boldsymbol{z})/R_i}\right) + 0.5 \text{ if } d_i(\boldsymbol{z}) \leq R_i \\ 0.5\left(\dfrac{1}{1 + \lambda_2(d_i(\boldsymbol{z}) - R_i)}\right) \text{ otherwise} \end{cases} \tag{10}$$

where parameters λ_1 and λ_2 are constants which satisfy $\lambda_2 = 1/(R_i(1 + \lambda_1))$.

It could be observed that the greater $d_i(\boldsymbol{z})$ is, the lower $\mu_i(\boldsymbol{z})$ is. The value of $\mu_i(\boldsymbol{z})$ is larger than 0.5 if $\boldsymbol{z}$ is inside the sphere ($d_i(\boldsymbol{z}) < R_i$). In [11], the sample $\boldsymbol{z}$ belongs to the class with the maximum criteria:

$$Class(\boldsymbol{z}) = arg \max_i \mu_i(\boldsymbol{z}) \tag{11}$$

Nevertheless, the capability of detection of novel cluster is still omitted in this decision rule. In this paper, in order to detect the potential novel cluster, the decision rule is modified by adding a judgment of the down limit of $\max \mu_i(\boldsymbol{z})$, as

$$Class(\boldsymbol{z}) = \begin{cases} \arg\max_i \mu_i(\boldsymbol{z}) & \text{if } \max \mu_i(\boldsymbol{z}) \geq Th_i \\ \text{novel} \quad \text{if } \max \mu_i(\boldsymbol{z}) < Th_i \end{cases} \tag{12}$$

where Th_i is the threshold of the class ω_i. The threshold is set based on the calibration of the validation dataset. For

[1]Throughout this paper, the Gaussian kernel $k(\boldsymbol{z}_n, \boldsymbol{z}_m) = \exp(-||\boldsymbol{z}_n - \boldsymbol{z}_m||^2/\sigma)$ will be used.

instance, say we have N_i validation data in class ω_i, Th_i is set using *3-sigma law*

$$Th_i = \bar{\mu_i}(\boldsymbol{z}) - 3\sqrt{\frac{1}{N_i - 1}\sum_{\boldsymbol{z}_n \in \omega_i} \left(\mu_i(\boldsymbol{z}_n) - \bar{\mu_i}(\boldsymbol{z})\right)^2} \quad (13)$$

where $\bar{\mu_i}(\boldsymbol{z}) = \sum_{\boldsymbol{z}_n \in \omega_i} \mu_i(\boldsymbol{z})/N_i$. Thus, it can be confirmed that only 0.15% of the samples in the validation dataset are mis-classified to the novel class.

III. Experimental results

A. PEMFC stack and testbench

A 5-kW PEMFC testbench is used to carry out the various experiments (see Fig. 3). Other than the fuel cell stack, the following subsystems are consisted in the testbench:

- Air supply subsystem: the flow rate, pressure, temperature and hygrometry level at the air inlet can be regulated to the required conditions.
- Hydrogen supply subsystem: the flow rate and pressure at the hydrogen inlet, temperature and hygrometry level at the hydrogen inlet can be regulated to the required conditions.
- Temperature subsystem: this subsystem is dedicated to the regulation of stack temperature.
- Electronic load: the load current can be flexibly varied through an electronic load.
- Control/supervision unit: The controls of the test bench and the parameter monitoring are fulfilled using National Instruments Materials and Labview softwares [13].

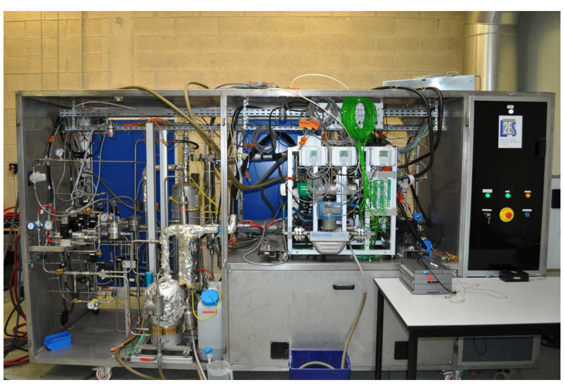

Fig. 3. Photograph of the used 5-kW PEMFC testbench

A 40-cell stack was used to carry out various experimental tests, including the ones under fault conditions. The stack was fabricated by french research organization CEA (Alternative Energies and Atomic Energy Commission) in the framework of the French ANR DIAPASON project. The nominal operating conditions of the stack are summarized in Table I.

B. Experimental database

The experiments of normal state and various fault states were carried out in the testbench. The concerned faults are listed in Table II. Each faulty experiment was repeated

TABLE I
NOMINAL CONDITIONS OF THE STACK

Parameter	Value
Stoichiometry H_2	1.5
Stoichiometry Air	2
Pressure at H_2 inlet	150 kPa
Pressure at Air inlet	150 kPa
Differential of anode pressure and cathode pressure	30 kPa
Temperature (exit of cooling circuit)	80 °C
Anode relative humidity	50%
Cathode relative humidity	50%
Current	110 A
Voltage per cell	0.7 V
Electrical power	3080 W

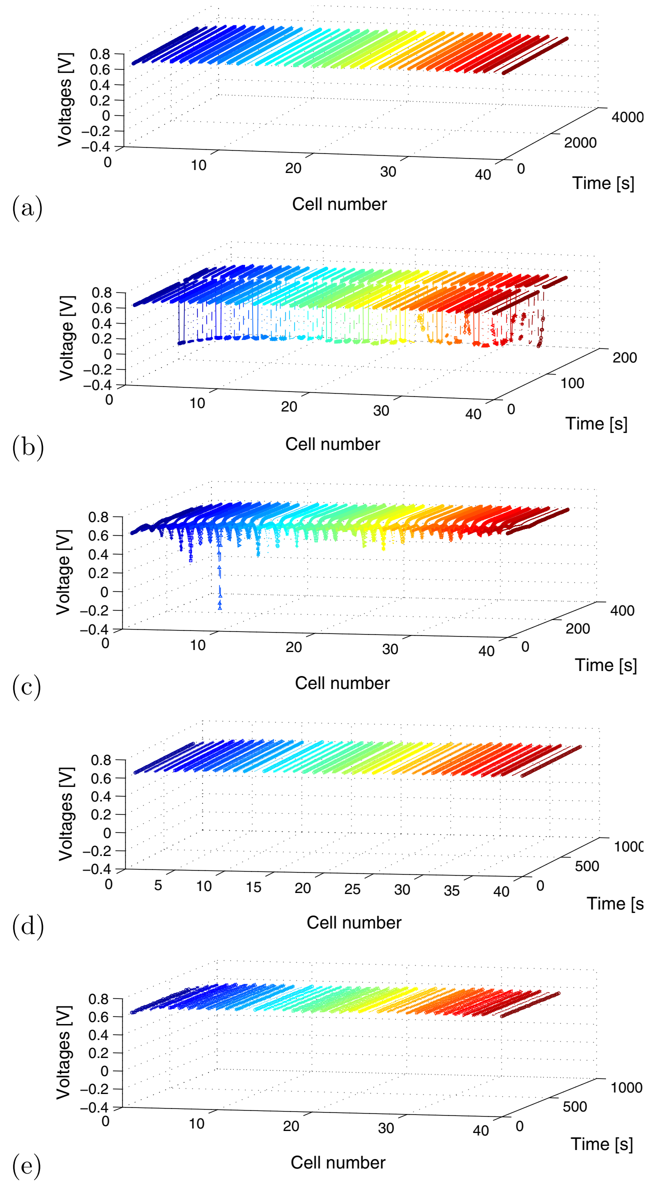

Fig. 4. Evolution of cell voltages in different states. (a) Nl, (b) F_1, (c) F_2, (d) F_3, (e) F_4

TABLE II

CONCERNED STATES (CLASSES)

Health state description	Location	Notation
Nominal operating state	Whole system	Nl
High current pulse or short circuit	Default of electric circuit	F_1
Stop cooling water	Default of temperature subsystem	F_2
High air stoichiometry	Default of air supply subsystem	F_3
Low air stoichiometry	Default of air supply subsystem	F_4

several times. Although various physical variables had been sampled, only the cell voltages sampled during the experiments were drawn to construct the training dataset and the test (validation) dataset. That is to say, the original dimensional number of the data is 40, which is equal to the fuel cell number. For each state, data from one (or several) experiment(s) were used as training data, while data from others were considered as test data. The sample numbers used for training and testing in different classes are summarized in Table III.

TABLE III

SAMPLE NUMBERS OF THE TRAINING AND TEST DATA

	Nl	F_1	F_2	F_3	F_4
Training data	2202	108	302	400	400
Test data	2200	31	201	2500	2200

The evolution of cell voltages in different states is shown in Fig. 4 (F_1, F_2 occurred in-between the experiments). It could be found that the voltages of different cells have the different responses to different faults. Actually, this character is serviceable for fault diagnosis.

C. Results and discussion

1) Train data from all states: To verify the proposed diagnosis strategy, firstly, we took all of the 5 states into consideration. After training FDA, the data of 5 states were projected to a 4-dimensional space. Fig. 5 shows the first 3 features of the *projecting vectors*. The samples from different classes are dispersive from visual point of view.

The extracted features were used to train the SSM-SVM classifier afterwards. By using the trained FDA and SSM-SVM models, the diagnostic accuracy of the test dataset was evaluated. The confusion matrix is shown in Table IV. From the table, it could be seen that the diagnostic accuracy is more than 95% for each class except the class F_4. Actually, F_4 is the most light fault which has some overlaps with the normal state. The phenomenon can also be observed from Fig. 5.

2) Leave one fault class as a novel class: To test the performance of novel failure mode detection, the data from one fault state was assumed as a novel class, while the models were trained using the data from other classes. The confusion matrices, as the classes F_1, F_2, F_3, F_4 are considered as novel classes, are shown respectively in Table V-VIII.

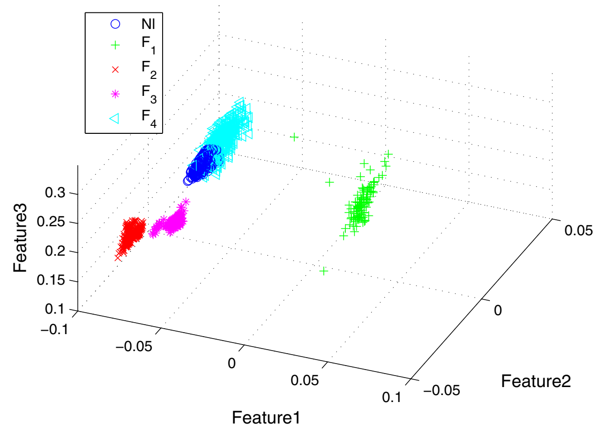

Fig. 5. Data of 5

TABLE IV

CONFUSION MATRIX (%) AFTER TRAINING ON ALL CLASSES

		Diagnosed class				
		Nl	F_1	F_2	F_3	F_4
Actual class	Nl	95.95	0.14	0	0	3.91
	F_1	3.23	96.77	0	0	0
	F_2	0	4.76	95.24	0	0
	F_3	0	2.44	0	97.56	0
	F_4	8.55	0.95	0	0	90.50

TABLE V

CONFUSION MATRIX (%) WITH F_1 AS NOVEL CLUSTER

		Diagnosed class				
		Nl	F_2	F_3	F_4	NC
Actual class	Nl	98.91	0	0	1.09	0
	F_2	0	95.24	0	0	4.76
	F_3	0	0	99.00	0	1
	F_4	6.77	0	0	93.23	0
	NC	3.23	0	0	0	96.77

TABLE VI

CONFUSION MATRIX (%) WITH F_2 AS NOVEL CLUSTER

		Diagnosed class				
		Nl	F_1	F_3	F_4	NC
Actual class	Nl	94.95	0	0	4.32	0.73
	F_1	3.32	83.87	0	0	12.9
	F_3	0	0	99.00	0	1.00
	F_4	8.55	0	0	90.36	1.09
	NC	0	0	0	0	1

TABLE VII

CONFUSION MATRIX (%) WITH F_3 AS NOVEL CLUSTER

		Diagnosed class				
		Nl	F_1	F_2	F_4	NC
Actual class	Nl	94.68	0	0	4.68	0.64
	F_1	3.23	90.32	0	0	6.45
	F_2	0	0	95.24	0	4.76
	F_4	9.73	0	0	89.45	0.82
	NC	0.12	0	0	4.52	95.36

TABLE VIII

CONFUSION MATRIX (%) WITH F_4 AS NOVEL CLUSTER

		Diagnosed class				
		Nl	F_1	F_2	F_3	NC
Actual class	Nl	99.59	0	0	0	0.41
	F_1	3.23	96.77	0	0	0
	F_2	0	0	95.24	0	4.76
	F_3	0	0	0	99.16	0.84
	NC	46.91	13.23	0	0	39.86

For all the cases, the probabilities, that the data located in the known classes are mis-classified to the novel classes, are generally low. For the cases that F_1, F_2, and F_3 are considered as novel classes, the probabilities of detection the novel class are more than 95%; while for the case that F_4 is considered as novel class, the probability is 39.86%, which is a low level. It can be deduced that, it is difficult to detect the data in the novel class, when they are closed to the known classes.

3) Real-time implementation: It should be emphasized that although it seems to be sophisticated to address the training of the FDA and SSM-SVM models (The eigenvalue problem and QP problem for instance), the performing procedures of the trained models are much less burdened. In our case, the approach can be performed within 0.7 ms using ordinary PC. Hence, it is possible to implement the approach in real-time.

IV. CONCLUSION AND PERSPECTIVE

A data-based diagnostic strategy is proposed for PEMFC system. In the strategy, features are extracted from cell voltages by using FDA, and the extracted features are diagnosed by using the classification method SSM-SVM. Moreover, the potential novel failure mode can be detected. The test results of a 40-cell PEMFC stack have shown that, the probabilities of detecting and isolating the known fault are high in most cases. And the novel fault type can also be detected with a high accuracy in most cases.

At present, the work of coding the approach in an embedded computing chip and online testing is in process. Additionally, how to online update the trained models is also our focus.

ACKNOWLEDGE

This Ph.D. work is a contribution to the ANR DIAPASON2 project (fuel cell diagnosis methods for vehicle and stationary applications 2^{nd} phase). The authors would like to thank the partners who supported us the experimental data.

REFERENCES

[1] D. Hissel, S. Member, D. Candusso, and F. Harel, “Fuzzy-Clustering Durability Diagnosis of Polymer Electrolyte Fuel Cells Dedicated to Transportation Applications,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 2414–2420, 2007.

[2] N. Yousfi Steiner, D. Hissel, P. Moçotéguy, and D. Candusso, “Diagnosis of polymer electrolyte fuel cells failure modes (flooding & drying out) by neural networks modeling,” *International Journal of Hydrogen Energy*, vol. 36, no. 4, pp. 3067–3075, Feb. 2011.

[3] N. Yousfi Steiner, D. Hissel, P. Moçotéguy, and D. Candusso, “Non intrusive diagnosis of polymer electrolyte fuel cells by wavelet packet transform,” *International Journal of Hydrogen Energy*, vol. 36, no. 1, pp. 740–746, Jan. 2011.

[4] J. Hua, J. Li, M. Ouyang, L. Lu, and L. Xu, “Proton exchange membrane fuel cell system diagnosis based on the multivariate statistical method,” *International Journal of Hydrogen Energy*, pp. 1–10, 2011.

[5] Z. Li, S. Giurgea, R. Outbib, and D. Hissel, “Online diagnosis of pemfc by combining support vector machine and fluidic model,” *Fuel Cells*, pp. n/a–n/a, 2014.

[6] P. E. H. Richard O. Duda and D. G. Stork, *Pettern Classification*, 2nd ed. John Wiley & Sons, 2001.

[7] Z. Li, R. Outbib, D. Hissel, and S. Giurgea, “Online Diagnosis of PEMFC by Analyzing Individual Cell Voltages,” in *2013 European Control Conference (ECC)*, Zurich, Switzerland, July 17-19 2013.

[8] L. Cao, K. Chua, W. Chong, H. Lee, and Q. Gu, “A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine,” *Neurocomputing*, vol. 55, no. 1, pp. 321 – 336, 2003.

[9] D. Olson and D. Delen, *Advanced Data Mining Techniques*. Springer, 2008.

[10] C.-w. Hsu and C.-j. Lin, “A Comparison of Methods for Multiclass Support Vector Machines,” *IEEE Transaction on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.

[11] P. Y. Hao and Y. H. Lin, “A new multi-class support vector machine with multi-sphere in the feature space,” in *Proceedings of the 20th international conference on Industrial, engineering, and other applications of applied intelligent systems*, ser. IEA/AIE’07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 756–765.

[12] P. Laskov, C. Gehl, S. Krüger, and K.-R. Müller, “Incremental support vector learning: Analysis, implementation and applications,” *The Journal of Machine Learning Research*, vol. 7, pp. 1909–1936, 2006.

[13] D. Candusso, F. Harel, a. Debernardinis, X. Francois, M. Pera, D. Hissel, P. Schott, G. Coquery, and J. Kauffmann, “Characterisation and modelling of a 5kW PEMFC for transportation applications,” *International Journal of Hydrogen Energy*, vol. 31, no. 8, pp. 1019–1030, Jul. 2006.