TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7664~7670
e-ISSN: 2087-278X

Fault Diagnosis for Fuel Cell based on Naive Bayesian Classification

Liping Fan*[1], Xing Huang[2], Liu Yi[3]
[1]College of Environment and Safety Engineering, Shenyang University of Chemical Technology, Shenyang, 110142, China
[2]College of Information Engineering, Shenyang University of Chemical Technology, Shenyang, 110142, China
[3]Branch Company of Rolling Equipment, North Heavy Industry Group Co., Ltd, Shenyang, Liaoning, 110141, China
*Corresponding author, e-mail: flpsd@163.com

Abstract

Many kinds of uncertain factors may exist in the process of fault diagnosis and affect diagnostic results. Bayesian network is one of the most effective theoretical models for uncertain knowledge expression and reasoning. The method of naive Bayesian classification is used in this paper in fault diagnosis of a proton exchange membrane fuel cell (PEMFC) system. Based on the model of PEMFC, fault data are obtained through simulation experiment, learning and training of the naive Bayesian classification are finished, and some testing samples are selected to validate this method. Simulation results demonstrate that the method is feasible.

Keywords: *proton exchange membrane fuel cell (PEMFC), fault diagnosis, naive bayesian classification*

1. Introduction

The world is facing an energy crisis as well as significant environmental problems. Fossil fuels such as petroleum, natural gas and coal are the main resources for generating electricity. But they also have been major contributors to environmental problems. Renewable bio energy is viewed as one of the ways to alleviate the current global warming crisis. Major efforts are devoted to developing alternative electricity production methods [1-3].

Fuel cells are promising energy sources which directly convert chemical energy of fuel into electrical energy through chemical reaction [4], with almost null pollutant emissions [5]. proton exchange membrane fuel cell (PEMFC), as the most popular kind of fuel cells in the residential and vehicular applications, when is compared with the other fuel cells, is able to efficiently generate high power densities [6]. Research and application of proton exchange membrane has attracted more and more attentions [7]. Current information shows that PEMFC has been applied in the fields of portable power supply, vehicle power supply, distributed power plants and some other fields [8].

The performance of PEMFC, being important and getting more and more attention in recent years, is known to be influenced by many parameters such as operating temperatures both fuel cell and humidifiers, pressure, flow rates and relative humidity of fuel and oxidant gases. The systems of fuel cell are required to have higher safety and stability. The energy generation systems based on fuel cells are complex since they involve thermal, fluidic and electrochemical phenomena [9]. Faults are events that cannot be ignored in fuel cell system, and their consideration is essential for improving the operability, flexibility, and autonomy of them [10, 11].

Bayesian network is one kind of graph model which describes the relation among data variables, and is also a model used to infer. It describes the correlation between data with probability measure weight, thus to solve the inconsistency between data, even the problems of mutual independence, can expediently deal with incomplete information problems. Bayesian network can be used in on-line and off-line fault diagnosis of a complex system [12, 13].

A classifier is a function that maps an instance into a class label. Traditional approaches to this problem include decision trees, neural networks, and classical statistical

Received June 30, 2013; Revised August 20, 2013; Accepted September 5, 2013

methods. More recently, Bayesian networks have also been successfully applied to classification problems in many ways by inducing classifiers using different types of Bayesian network learning algorithms [14]. Bayesian network classifier is a kind of Bayesian network that used to complete the task of classification. This paper uses the method of naive Bayesian classifier in the fault diagnosis of the system of proton exchange membrane fuel cell. Different operating conditions which have influences on the PEMFC performance are considered, and the corresponding models of fault diagnosis are established.

2. Model of PEMFC

As a tool for the optimization of the design of the fuel cell, the mathematical model and simulation is necessary. It is important to have an appropriate model to estimate the overall performance of proton exchange membrane fuel cell operating conditions [15].

Fuel cells are electrochemical devices that convert the chemical energy of a fuel (hydrogen) and an oxidizer (oxygen) directly into electricity. The fuel cell consists of an electrolyte between two electrodes. The electrolyte allows the positive ions (protons) to pass through while blocking the electrons. Hydrogen gas passes over one electrode (anode), and with the help of a catalyst, separates into electrons and hydrogen protons. A graphic representation of a PEM fuel cell structure is given in Figure 1 [16].

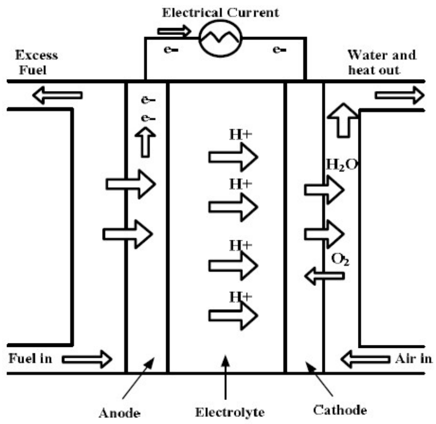

Figure 1. Schematic Diagram of PEMFC

PEM fuel cell electrochemical process starts on the anode side where H_2 molecules are brought by flow plate channels. Anode catalyst divides hydrogen on protons H^+ that travel to cathode through membrane and electrons e^- that travel to cathode over external electrical circuit. At the cathode hydrogen protons H^+ and electrons e^- combine with oxygen O_2 by use of catalyst, to form water H_2O and heat. Described reactions can be expressed by the following equations [17, 18]:

$$\mathrm{H_2 \rightarrow 2H^+ + 2e^-} \ \text{(Anode)} \tag{1}$$

$$\mathrm{\frac{1}{2}O_2 + 2H^+ + 2e^- \rightarrow H_2O} \ \text{(Cathode)} \tag{2}$$

$$\mathrm{H_2 + \frac{1}{2}O_2 \rightarrow H_2O + heat{+}electricity} \ \text{(Total)} \tag{3}$$

The output voltage V_{fc} of a single cell can be demonstrated by the following equations:

$$V_{\mathrm{fc}} = E_{\mathrm{nernst}} - V_{\mathrm{act}} - V_{\mathrm{ohmic}} - V_{\mathrm{con}} \tag{4}$$

The output power of the single fuel cell is:

$$E_{\mathrm{nernst}} = 1.229 - 0.85\times10^{-3}(T_{\mathrm{fc}} - 298.15) + 4.31\times10^{-5} T_{\mathrm{fc}}\left[\ln(P_{H_2}) + \frac{1}{2}\ln(P_{O_2})\right] \quad (5)$$

Where i is the output current of the fuel cell. A generally accepted dynamic model of PEMFC is shown in Figure 1, where q_{O2} is the input molar flow of oxygen, q_{H2} is he input molar flow of hydrogen, K_{H2} is the hydrogen valve molar constant, and K_{O2} is the oxygen valve molar constant, P_{H2} and P_{O2} are pressures of hydrogen and oxygen; V_{fc}^{*} is the given output voltage, V_{fc} denotes the output voltage of PEMFC, E_{nernst} is the thermodynamic potential of the cell representing its reversible voltage, V_{ohmic} is the ohmic voltage drop associated with the conduction of protons through the solid electrolyte and electrons through the internal electronic resistance, V_{con} represents the voltage drop resulting from the mass transportation effects, V_{act} is the voltage drop due to the activation of the anode and the cathode [19, 20].

An accepted dynamic model of the PEMFC is shown in Figure 2. According to the above described mathematical model, a Matlab/Simulink simulation model of PEMFC can be set up [21]. Parameters used in the simulations are those of the Ballard Mark V fuel cell [22].

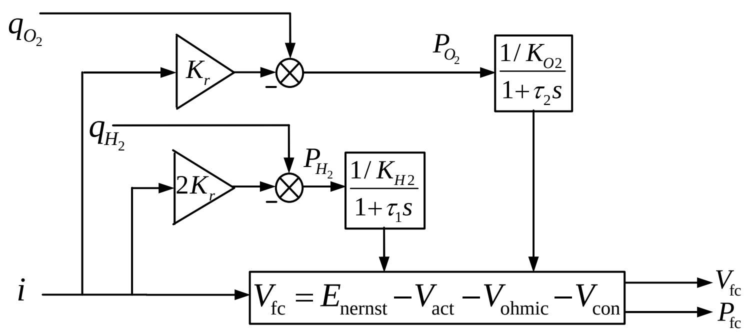

Figure 2. PEMFC Dynamic Model

3. Naive Bayesian Classifier

Many tasks, including fault diagnosis, pattern recognition and forecasting, can be viewed as classification. Classification is the process of using a model to predict unknown values (output variables), using a number of known values (input variables).

In order to perform classification, first we need to model the relationship between the input variables and the output variables we are predicting. This process involves learning a model using data in which both the input variables and the output variables are present. Expert opinion can also be used to build or enhance a model. This model can subsequently be used on unseen data in which only the input data is present, in order to predict the output variables.

Bayesian networks are widely used to perform classification tasks, with the following advantages:

a. Based on probability theory.
b. Not a black box approach.
c. Allows rich structure.
d. Can mix expert opinion and data to build models.
e. Backwards reasoning-in addition to predicting outputs given inputs, we can use output values to infer inputs.
f. Support for missing data during learning and classification.

Bayesian network classifier is a typical classification model based on statistical method. In Bayesian network classifier, the priori probability of the events is cleverly linked to the posterior probability, using priori information and sample data to determine the posterior probability of events [23].

Suppose $U = \{X_1, X_2, \ldots, X_n, C\}$ is a finite set of discrete random variables, where $X_1, X_2, \ldots, X_n$ are attribute variables, C is a class variable and its value range is $\{c_1, c_2, \ldots, c_m\}$. If x_i is the value of X_i. then the probability of case $I_i=(x_1, x_2, \ldots x_n,)$ belonging to c_j can be expressed by Bayes' theorem as:

$$P\left(c_j \mid x_1,x_2,\cdots,x_n\right)=\frac{P\left(x_1,x_2,\cdots,x_n \mid c_j\right)\cdot P\left(c_j\right)}{P\left(x_1,x_2,\cdots,x_n\right)}=\alpha P\left(c_j\right)\cdot P\left(x_1,x_2,\cdots,x_n \mid c_j\right) \quad (6)$$

Where α is the regularization factor, $P\left(c_j\right)$ is a prior probability of class c_j, $P\left(c_j \mid x_1,x_2,\cdots,x_n\right)$ is a posterior probability of class c_j which reflects the influence of sample data to class c_j. Equation (6) can also be expressed as:

$$P\left(c_j \mid x_1,x_2,\cdots,\ x_n\right)=\alpha P\left(c_j\right)\cdot\prod_{i=1}^{n}P\left(x_i \mid x_1,x_2,\cdots,x_{i-1},c_j\right) \quad (7)$$

According to the criterion of Bayesian maximum a posteriori (MAP) estimation , for a certain given case $I_i=(x_1, x_2, \ldots, x_n)$, Bayesian network classifiers will choose the class c_j which can make the posterior probability $P\left(c_j \mid x_1,x_2,\cdots,x_n\right)$ maximum as the class label of case I_i. Therefore, the key of Bayesian classifier is how to calculate $P\left(x_i \mid x_1,x_2,\cdots,x_{i-1},c_j\right)$. The differences between all kinds of Bayesian classifiers lie in the different calculation methods for $P\left(x_i \mid x_1,x_2,\cdots,x_{i-1},c_j\right)$ [24].

Bayesian networks are powerful tools for knowledge representation and inference under conditions of uncertainty, but they were not considered as classifiers until the discovery that Naïve-Bayes, a very simple kind of Bayesian networks that assumes the attributes are independent given the class node, are surprisingly effective [25].

Naive Bayesian is one of the most effective and efficient classification algorithms. The naive Bayesian classifier is a straightforward and widely used method for supervised learning. It is one of the fastest learning algorithms, and can deal with any number of features or classes. It has two advantages over many other classifiers. First, it is easy to construct. Second, the classification process is very efficient. Despite of its simplicity in model, naive Bayesian performs surprisingly well in a variety of problems. Furthermore, naive Bayesian learning is robust enough that small amount of noise does not perturb the results.

An advantage of the naive bayes classifier is that it only requires a small amount of training data to estimate the parameters (means and variances of the variables) necessary for classification. Because independent variables are assumed, only the variances of the variables for each class need to be determined and not the entire covariance matrix. Naive Bayesian classifier assumed that all input attributes are conditionally independent. Consequently, Equation (2) can be expressed as:

$$P\left(c_j \mid x_1,x_2,\cdots,x_n\right)=\alpha P\left(c_j\right)\cdot\prod_{i=1}^{n}P\left(x_i \mid c_j\right) \quad (8)$$

Although the assumption restricts the applicability of naive Bayesian classifier to a certain extent, however many researches show that in the practical applications, naive Bayesian network also showed considerable robustness and efficiency.

4. Fault Diagnosis of PEMFC

In order to construct a feasible fault diagnosis scheme for PEMFC based on naive Bayesian classifier, the following steps can be referred:

1) Choose the appropriate model of Bayesian network classifier;
2) By acquisition and classification of PEMFC examples, determine the attribute variables and fault class variables;

3) Learning the parameter of naive Bayesian classifier by utilizing expert knowledge or through the training sample;
4) Fault diagnosis analysis by Bayesian classifier.

According to the model of PEMFC system we established in our previous work, the influences of operating temperature and input hydrogen and oxygen pressure on the output performance of proton exchange membrane fuel cell system are analyzed. The input hydrogen pressure and the operating temperature of PEMFC system are used as fault class variables of the diagnosis system and the output voltage and output power of PEMFC system are used as attribute variables. Pulse signal is used as interference. Fault data used as training sample are acquired through the model simulation. Three kinds of training samples which are different in size are adopted, one for 375 fault data, one for 750 fault data and the other for 1500 fault data. Under each group of training samples, five testing samples are selected to test the results of PEMFC system fault diagnosis based on this method.

The Naïve bayes classifier learns from training data the conditional probability of each variable X_i given the class label C. Figure 3 is the model of naive Bayesian network used in this paper, in which X means the output parameters of the PEMFC system, C_1 denotes the module fault of working temperature in PEMFC system, C_2 means the module fault of the input hydrogen pressure, C_3 means the module fault of the input oxygen pressure.

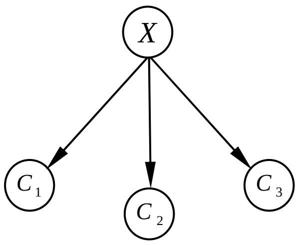

Figure 3. Naive Bayesian model for fault diagnosis of PEMFC system

Table1. Part of the Training Samples

N	Fault category	X_1(Power)	X_2(Voltage)
1	C_1	0.25139340035598	1.5855390261863
2	C_2	0.25132235202399	1.5853149593188
3	C_3	0.25369373645009	1.5927766210304
4	C_1	0.25139340035597	1.5855390261863
5	C_2	0.25149965223106	1.5858740562575
6	C_3	0.25427729351951	1.5946074548914
7	C_1	0.25139340035596	1.5855390261862
8	C_2	0.25168052818246	1.5864442258789
9	C_3	0.25484560424687	1.5963884372134

Bayesian classification is to calculate the sample's posterior probability belongs to a kind of class through the existing evidence. The class which has the maximum posteriori probability is the one that the sample belongs to. Therefore, Bayesian network classifiers do not need large training samples, but it can get higher accuracy of fault diagnosis.

Parameters of Bayesian network classifier are generally divided into two parts, which are called structural parameters and probability parameters. However, the naive Bayesian classifier doesn't need to learn the structural parameters, so here only the probability parameters are discussed.

The parameters can be determined according to the expert experience, or learned from the training samples. In this paper, the latter way is adoptted. The corresponding probability parameters are gained by learning the training samples. Three kinds of training samples in different sizes are selected. Five testing samples are selected for each kind of the training samples to compare and analysis the diagnosis results. Parts of the training samples are shown in Table 1. The specific diagnosis results for the PEMFC system mentioned above are shown in Figure 4 to Figure 6.

Three kinds of training samples are used to naive Bayesian classifier for PEMFC system fault diagnosis. Accuracy comparison between the three kinds of diagnosis mode is shown in Table 2. From these results we can see that with the expansion of the training sample, the accuracy of fault diagnosis is raising. Even if the fault data is 375 training samples, the accuracy of fault diagnosis can reach above 80%. The average accuracy of five kinds of testing samples is about 83.42%. When the training samples expand to 750, the accuracy of fault diagnosis can reach above 85%, and the average accuracy of five kinds of testing sample under these training samples is about 86.83%. When the training samples expand to 1500, the accuracy of fault diagnosis is over 90%, the average accuracy of five kinds of testing sample under these training samples is about 91.50%. Judging from the experimental results, application of naive Bayesian classifier to fault diagnosis for PEMFC system is feasible.

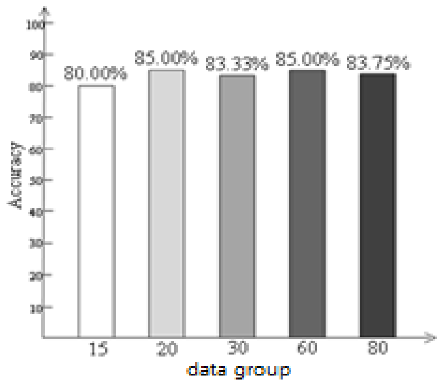

Figure 4. Diagnosis Result Corresponding to 375 Training Samples

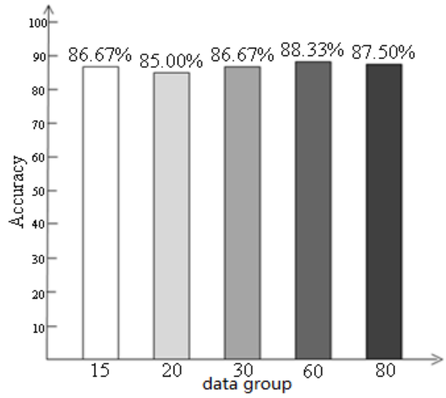

Figure 5. Diagnosis Result Corresponding to 750 Training Samples

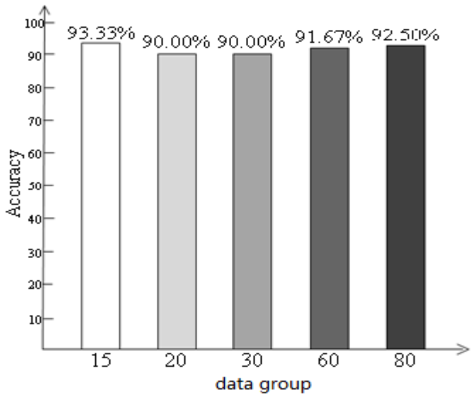

Figure 6. Diagnosis Result Corresponding to 1500 Training Samples

Table 2. Comparison Results of Accuracy between Three Kinds of Training Way

Testing sample	Training sample 375	750	1500
15	80.80%	86.67%	93.33%
20	85.00%	85.00%	90.00%
30	83.33%	86.67%	90.00%
60	85.00%	88.33%	91.67%
80	83.75%	87.50%	92.50%

5. Conclusion

Fault diagnosis of PEMFC system based on the naive Bayesian classifier is presented and proved. A fault diagnosis method based on naive Bayesian classifier does not need a lot of prior data, but it can give a higher accuracy of fault diagnosis. Fault diagnosis based on naive Bayesian classifier for PEMFC system is feasible.

Acknowledgement

This work was supported the National Key Technology Research and Development Program of China under Grant 2012BAF09B01, and the Science and Technology Research Project of Liaoning Education Department of China under Grant L2012140.

References

[1] Das S, Mangwani N. Recent Development of Microbial Fuel Cell: a Review. *Journal of Scientific & Industrial Research*. 2010; 69(10): 727-731.
[2] Franks A, Nevin K. Microbial Fuel Cells, A Current Review. *Energies*. 2010; 3(5): 899-919.
[3] Logan B. Scaling up Microbial Fuel Cells and Other Bioelectrochemical Systems. *Applied Microbiology and Biotechnology*. 2010; 85(6): 1665-1671.

[4] Gencoglu M, Ural Z. Design of a PEM Fuel Cell System for Residential Application. *Int J Hydrogen Energ.* 2009; 34(12): 5242-5248.
[5] Samosir A, Sutikno T, Yatim A. Dynamic Evolution Control for Fuel Cell DC-DC Converter. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2011; 9(1): 183-190.
[6] Kunusch C, Puleston P, Mayosky M, Riera J. Sliding Mode Strategy for PEM Fuel Cells Stacks Breathing Control Using a Super-Twisting Algorithm. *IEEE Transactions on Control Systems Technology*. 2009; 17(1): 167-174.
[7] Atia DM, Fahmy FH, Ahmed NM, Dorrah HT. A New Control and Design of PEM Fuel Cell Powered Air Diffused Aeration System. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2012; 10(2): 291-302.
[8] Oszcipok M, Zedda M, Hesselmann J, Huppmann M, Wodrich M, Junghardt M, Hebling C. Portable Proton Exchange Membrane Fuel-Cell Systems for Outdoor Applications. *Journal of Power Sources.* 2006; 157(2): 666-673.
[9] Escobet T, Feroldi D, Puig V, Quevedo J, Rierab J, Serra M. Model-based Fault Diagnosis in PEM Fuel Cell Systems. *Journal of Power Sources*. 2009; 192(1): 216-223.
[10] Riascos L, Simoes M, Miyag P. On-line Fault Diagnostic System for Proton Exchange Membrane Fuel Cells. *Journal of Power Sources*. 2008; 175(1): 419–429.
[11] Quan R, Quan S, Huang L, Xie C, Chen Q. Information Fusion in Fault Diagnosis for Automotive Fuel Cell System Based on D-S Evidence Theory. *Journal of Computational Information Systems.* 2011; 7(1): 97-105.
[12] Shi Q, Liang S, Fei W, Shi Y, Shi R. Study on Bayesian Network Parameters Learning of Power System Component Fault Diagnosis Based on Particle Swarm Optimization. *International Journal of Smart Grid and Clean Energy*. 2013; 2(1): 132-137.
[13] Ding J, Kramer B, Bai Y, Chen H. Backward Inference in Bayesian Networks for Distributed Systems Management. *Journal of Network and Systems Management*. 2005; 13(4): 409-427.
[14] Acid S, Campos L, Castellano J. Learning Bayesian Network Classifiers: Searching in a Space of Partially Directed Acyclic Graphs. *Machine Learning*. 2005; 59(3): 213-235.
[15] Carnes B, Djilal N. Systematic Parameter Estimation for PEM Fuel Cell Models. *Journal of Power Sources*. 2005; 144(1): 83–93.
[16] Abdous F. Fuel Cell/DC-DC Convertor Control by Sliding Mode Method. *World Academy of Science, Engineering and Technology*. 2009; 25: 1012–1017.
[17] Moreira M, Silva G. A Practical Model for Evaluating the Performance of Proton Exchange Membrane Fuel Cells. *Renewable Energy*. 2009; 34(7): 1734–1741.
[18] Youssef M, Nadi K, Khalil M. Lumped Model for Proton Exchange Membrane Fuel Cell (PEMFC). *International Journal of Electrochemical Science*. 2010; 5: 267–277.
[19] Mammar K, Chaker A. Fuzzy Logic Control of Fuel Cell System for Residential Power Generation. *Journal of Electrical Engineering.* 2009; 60(6): 328-334.
[20] Rezazadeh A, Askarzadeh A, Sedighizadeh M. Adaptive Inverse Control of Proton Exchange Membrane Fuel Cell Using RBF Neural Network. *International Journal of Electrochemical Science*. 2011; 6: 3105-3117.
[21] Fan L. Simulation Study on the Influence Factors of Generated Output of Proton Exchange Membrane Fuel Cell. *Applied Mechanics and Materials*. 2012; 121-126: 2887-2891.
[22] Correa J, Farret F, Canha L, Simoe M. An Electrochemical-Based Fuel-Cell Model Suitable for Electrical Engineering Automation Approach. *IEEE Transaction on Industrial Electronics*. 2004; 51(5): 1103-1112.
[23] Chhillar S, Khehra B. Method for Forecasting the Volume of Applications Using Probabilistic Classifier Based on Bayesian Theorem For Recruitment in the Government Organizations. *International Journal on Computer Science and Engineering.* 2012; 4(12): 1915-1919.
[24] Zeng Y, Zhu J, Gong J. Fault Diagnosis on Cooling System of Ship Diesel Engine Based on Bayes Network Classifier. *Journal of Central South University (Science and Technology)* (in China). 2010; 41(4): 1379-1384.
[25] Friedman N. Bayesian Network Classifiers. *Machine Learning*. 1997; 29: 131-163.