Voltage Control of PEMFC Using A New Controller Based on Reinforcement Learning

M. Karimi, M. Imanzadeh, P. Farhadi, and N. Ghadimi

Abstract—**In this paper, one of the most important fuel cell membrane with proton exchange visitors (PEMFC) was introduced, implemented and its voltage is controlled during load variations in a study system. Control algorithm used in this paper for the fuel cell plant during the variation of time for the input current to the membrane (*Ist*) is one of the algorithms which is based on reinforcement learning retention (RL). To control the study plant, this algorithm to control the studied plant is trying to change the gain and pole and zero coefficients so as to prevent as far as possible the output voltage changes, the output voltage, and to keep constant Fuel Cell in a specific required amount. Simulation results verify good accuracy performance of the proposed controller to achieve the goals.**

Index Terms—**Lag-lead control, PEMFC, reinforcement learning, voltage control.**

I. INTRODUCTION

Proton Exchange Membrane Fuel Cells (*PEMFCs*) consist of an anode and cathode blades as well as an electrolyte between these blades. Hydrogen gas (H_2) that is achieved from methanol (CH_3OH) is entered to the bottom anode blade (negative electrode) and oxygen or air to the end of the other blade "(positive electrode cell) [1]. To generate electric energy from fuel cell, it is necessary to have fixed values for the output voltage of cell for various loads so as to supply the given demand. However, fuel cell output voltage is changed in different loads. In order to keep constant cell's voltage, it is vital to use a controller. For this, the most simple type controller that can be used is a PID or Lead-Lag controllers.

In [1], a type of fuzzy controller to control the output voltage of fuel cell is proposed. In order to control the voltage and current of Fuel Cell, in [2], networks of BP and RBF were used. Speed and accuracy of proposed algorithms in which is satisfactory for the system. In [3], artificial neural networks are used to control fuel cell temperature. Ref. [4] has been used to achieve good control and efficient parameter optimization of a neural controller. Fuel Cells are in type of multiple fuel cells, but it is assumed that the mass cathode and anode is compacted in one compact fuel cell [5].

Each proposed methods is used to control one of the parameters of a fuel cell in which the method of Fuzzy or neural network is used. Some of these systems, first, identify and then control the system which in turn leads to time-consuming control application and sometimes causing long transient response.

In this paper a simple Lead-Lag Controller is used to control Fuel Cell voltage. The difference in this control is that zero, pole and gain of controller are changed online based on learning algorithm called Reinforcement Learning. Some of the benefits of the proposed control may be as: simple controller, being resistant against load changes, includes good control characteristics, fast transient response, and zero steady state error.

II. FUEL CELL MODEL

In order to study the dynamic model of Fuel Cell, schematic and structure as well as performance of fuel cell must be considered. A fuel cell schematic system studied in this paper is shown in Fig. 1. The mass of the anode and cathode is considered as a compact anode and cathode [5].

In this paper, dynamic model of fuel cell is same as in [1]. Fuel cell output voltage is obtained by reducing voltage drop of the backward voltage. Eq. (1) shows the calculation of fuel cell output voltage [6-7].

$$V_s = n(E_{reversable} - V_{act} - V_{ohmic} - V_{con}) \tag{1}$$

where V_s is the fuel cell output voltage accumulation in *volts*, n is the number of existing cells compacted in the fuel cell, V_{act} is voltage drop resulting from anode and cathode activity in *volts*, V_{ohmic} is ohmic voltage drop in *volts* which is a certain amount of resistance from the transfer of electrons and protons in the electrolyte between the anode and cathode. V_{con} is a voltage caused by the mass transfer of oxygen and hydrogen in *volts*. In Eq. (1), $E_{reversable}$ is calculated by the following equation [1 and 8],

$$\begin{aligned} E_{reversable} &= 1.229 - 0.85\times 10^{-3}(T - 298.15) \\ &+ 4.3085\times T \times [ln(PH_2 + 0.5 ln(PO_2)] \end{aligned} \tag{2}$$

where T is the temperature of cells in *Kelvin*, PO_2 and PH_2 are effective partial pressures of oxygen and hydrogen gases in *atmospheres*, respectively, which can be calculated by following equations,

$$PO_2 = P_c - P_{H_2o}^{sat} - P_{N_2}^{channel} exp\left(\frac{0.291\left(\frac{i}{A}\right)}{T^{0.932}} \right) \tag{3}$$

Manuscript received May 18, 2012; revised June 23, 2012.

M. Karimi, M. Imanzadeh and P. Farhadi are with Young Researchers Club, Parsabad Moghan Branch, Islamic Azad University, Parsabad Moghan, Iran (emails: m.karimi@iaupmogan.ac.ir, m.imanzadeh@iaupmogan.ac.ir, pfarhadi@iaupmogan.ac.ir).

N. Ghadimi is with Electrical Engineering department, Ardabil Branch, Islamiz Azad Univeristy, Ardabil, Iran (e-mail: n.ghadimi@iau-ardabil.ac.ir).

DOI: 10.7763/IJIEE.2012.V2.200

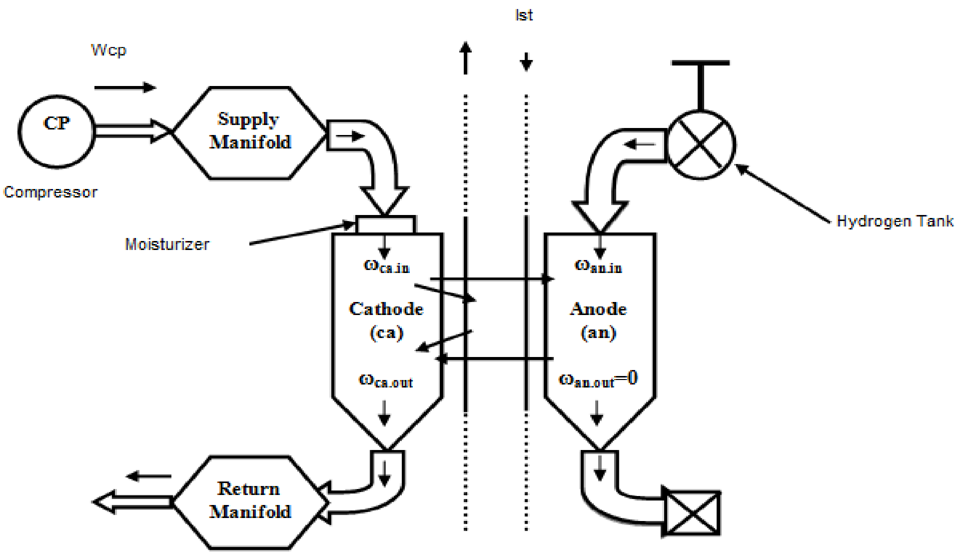

Fig. 1. Simplified Fuel Cell reactant supply system

$$PH_2 = 0.5P_{H_2o}^{sat}\left[\frac{1}{exp\left(\frac{1.635\left(\frac{i}{A}\right)}{T^{1.334}}\right).\left(\frac{P_{H_2o}^{sat}}{P_a}\right)} - 1\right] \quad (4)$$

where Pa, Pc are the anode and cathode inlet pressures in *atmospheres*, A is the effective electrode area in cm^2, i is the current of each cell in *amperes*, PHO is the amount of saturated water vapor pressure that its value is depend on the fuel cell. P_{N2} is the partial pressure of N_2 in the cathode gas flow channel in *atmospheres* obtained from the following Eq. (5) [9].

$$P_{N_2}^{channel} = \frac{0.79}{0.21}PO_2 \quad (5)$$

All parameters' values used in this paper are available in [1].

III. REINFORCEMENT LEARNING ALGORITHM AND THE CONTROL SYSTEM

Reinforcement learning *(RL)* method is based on trial and error algorithm which is trained to apply suitable control action on different plants indirectly. For each states of system in this method, one different control action is applied. For example if x_t and u_t denote the system state and system control at instant t, the state at instant $t+1$ is given by [5],

$$x_{t+1} = f(x_t, u_t) \quad u_t \in U, \forall t > 0 \quad (6)$$

For each control action in any intervals, remuneration or a penalty is dedicated. In other word, if this action improves the system error, remuneration will be considered if not a penalty. According to Fig. 1, the selection of remuneration or penalty is specified by a comparison between error at t and $t+1$. So remuneration selection at t for each action is performed using Eq. (7),

$$R(x_0, u\{t\}) = \sum_{t=0} \gamma^t r(x_t, u_t) \quad (7)$$

where in Eq. (7), Y is discount coefficient between 0 and 1, $r(x, u(t)) < B$ is a function of remuneration. Then from the obtained values a value function is defined from which control signal for the next time is selected.

$$V(x) = max_{u(t)} R(x, u(t)) \quad (8)$$

Using equation Bellman in [11], the value function can be defined as,

$$V(x) = max_{u \in U}[r(x,u) + \gamma V(f(x,u))] \quad (9)$$

The optimal control operation can also be obtained by Eq.(10).

$$u^*(x) = arg\ max_{u \in U}[r(x,u) + \gamma V(f(x,u))] \quad (10)$$

According to the mentioned equations, for each stage remuneration or penalty, function Q for each is expressed as a Eq. (11), and to select next action the value of this function would be determined.

$$Q(x,u) = r(x,u) + \gamma V(f(x,u)) \quad (11)$$

V(x) in Eq. (11) can be described in terms of Eq. (12).

$$V(x) = max_{u \in U} Q(x,u) \quad (12)$$

Thus the optimum control behavior is as Eq. (13).

$$u^*(x) = arg\ max_{u \in U} Q(x,u) \quad (13)$$

In the Lead-Lag control based on the mentioned method its function is shown in Eq. (14), for each gain, zero and pole of control, five different actions is placed. In each step, adaptive unit must implement actions from among 125 possible actions to select mode. In this process, some actions are eliminated through penalty turns and a random process to select another action between the remaining actions is carried out. The random selection with time is less to select a specific action with regard to favorite operation. Degree of freedom in this method is the number of matching parameters and the number of operations. The proposed control structure based on the method of reinforcing learning retention is shown in Fig. (2).

$$G_c(s) = k_p \frac{s+z}{s+p} \tag{14}$$

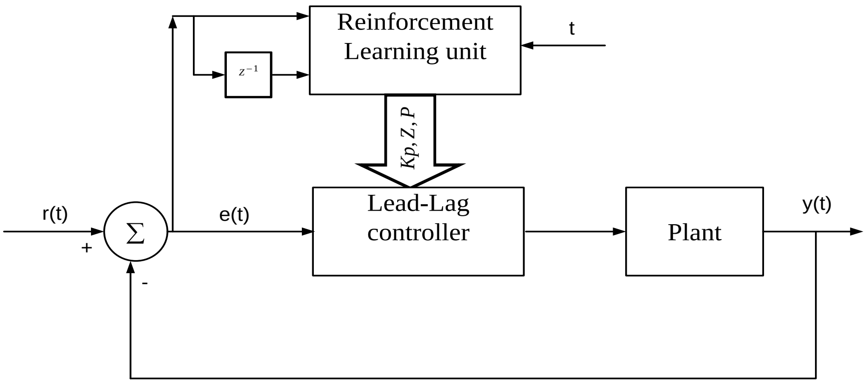

Fig. 2. Proposed control system.

IV. VOLTAGE CONTROL AGAINST VARIATIONS ON PEMFC (I_{ST})

By dynamically analysis of Fuel Cell, it is clear that its output voltage is varying sharply with output load changes; therefore, to use of fuel cell for energy generation, obviously the output voltage should be controlled. In order to control the output voltage, different controllers can be used among which the simplest type is the well-known classical *PID* controller.

In this section, output voltage is controlled in terms of load variations using classical *PID* controller that the coefficients are as (15). Fig. 3 shows anode and cathode pressures and the system load as well as output and reference voltages. For better illustration, the output voltage is shown in Fig. 4. According to this figure, it can be seen that in sharp changing areas that are system load, voltage-controlled characteristic is associated with an overshoot. Fig. 5 shows output voltage error (voltage difference to the reference voltage) that in its highest value almost equal to 1.1 *volt* causing the error to be 5.5% in these regions. It is worth to note that the amount of error in these areas is acceptable.

$$k_p = 10, k_i = 7, k_d = 0.4 \tag{15}$$

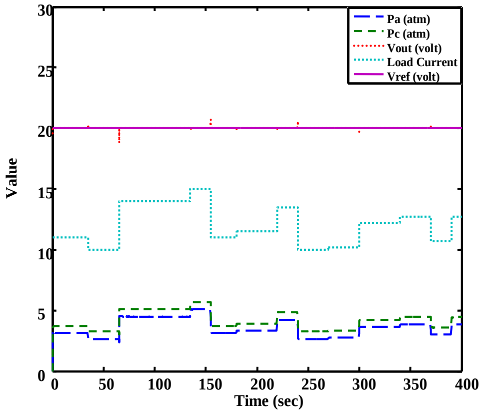

Fig. 3. The anode and cathode pressures and the system load as well as output and reference voltages related to the PID controller

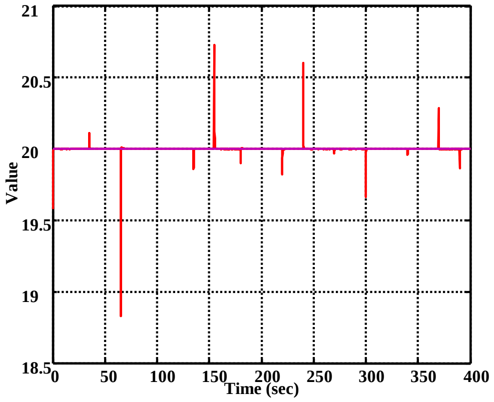

Fig. 4. Fuel cell output voltage related to PID controller

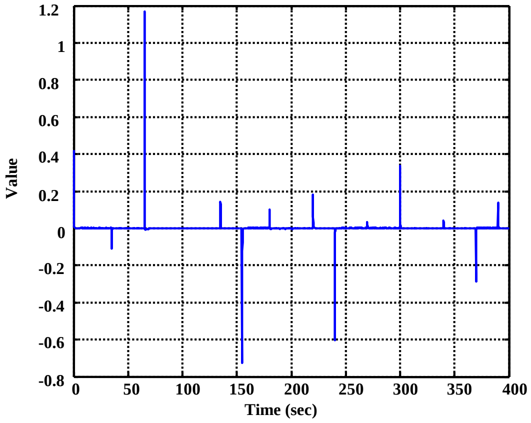

Fig. 5. The difference between output voltage and reference voltage related to the PID controller

Output results using the proposed controller with the expressed initial conditions in (16), is shown in Fig. 6. For better showing results, the proposed controller to the output voltage waveform is shown in Fig. 7. According to this figure it is clear that the maximum peak in the mode is 0.8, so respect to the previous mode (controller PID) this peak is much reduced. Other peak value is also reduced. Error related to the output voltage is shown in Fig. 8. Much reduction of error is the significant reason to use the proposed controller-based learning method.

$$k_{p0} = 400, P_0 = 2.9, Z_0 = 8.5 \qquad (16)$$

Variation of proportional coefficients, zero and pole controller based on the proposed method for reinforcing learning retention control system changes are seen in Fig. (9) to (11).

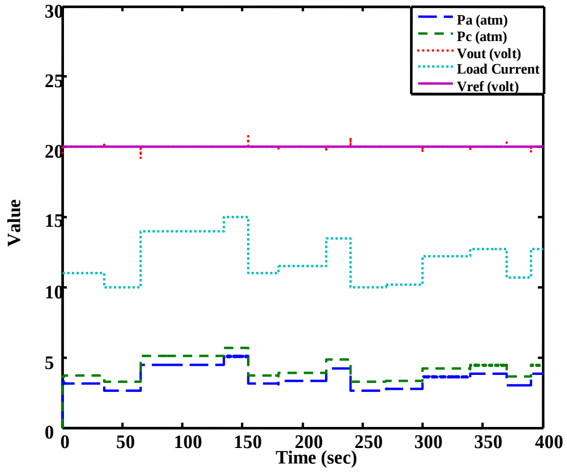

Fig. 6. Anode and cathode pressures and the system load as well as output and reference voltages related to the proposed controller

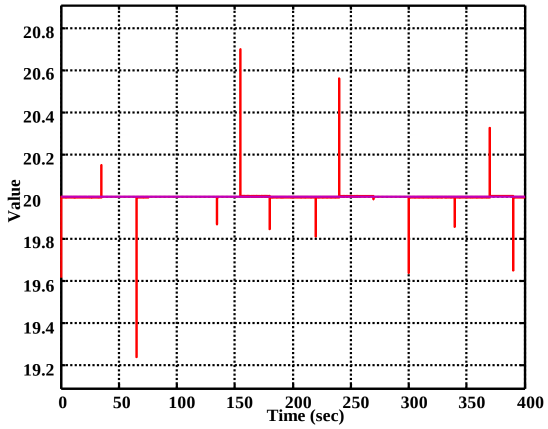

Fig. 7. Fuel cell output voltage related to the proposed controller

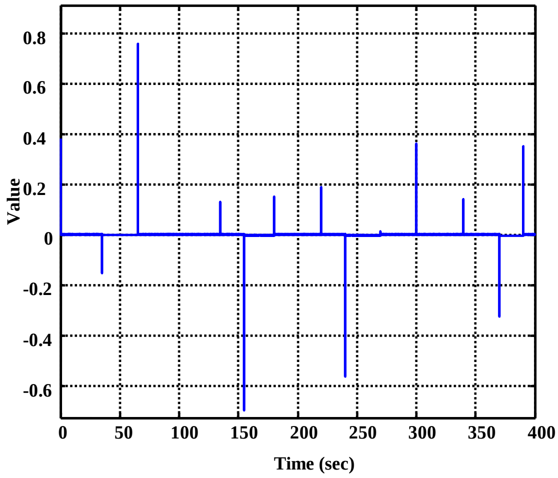

Fig. 8. Output voltage difference to the reference voltage to the controller proposed

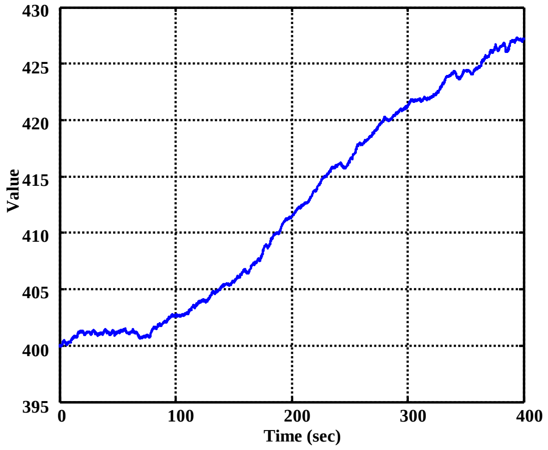

Fig. 9. Changes in interest ratio of proposed control method

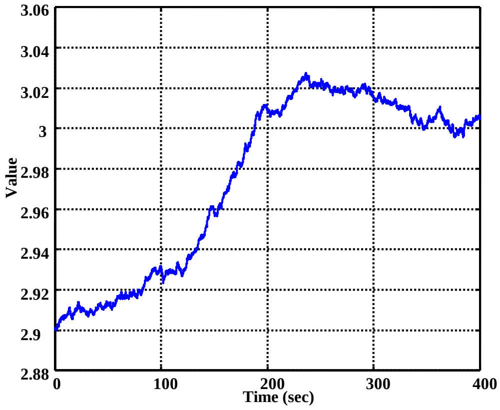

Fig .10. Changes in control center in the proposed control method

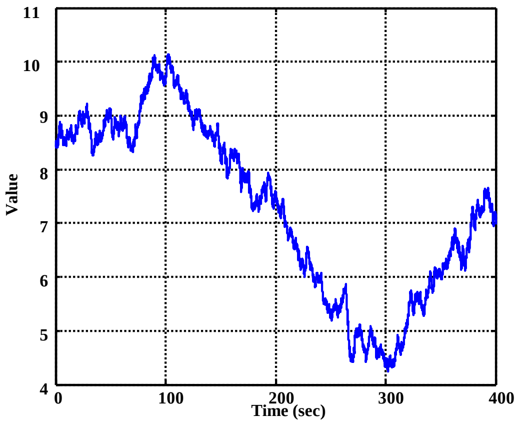

Fig .11. Zero change in control of proposed method controller

V. CONCLUSION

This paper, introduced a special Lead-Lag control method to control the PEMFC output voltage. In this method, coefficient, zero and pole of controller are changing according to changes based on reinforcing learning and to change their value every moment of time which Lag-Lead controller with the best values in its response to changes in input data and make optimal control on the study plant. Study plant in this paper is consisted of important plants in energy production in the world. High speed response in the used method, reducing errors and lack of major distortions which usually seen in adaptive control methods are the main reasons for using this control method.

REFERENCES

[1] M. Zhigun, Z. Xinjian, and C. Guangyi “Design and Simulation of Fuzzy Controller for PEMFCs,” *IEEE Conference*, 2005, pp. 220-224.

[2] A. Saengrung, A. Abtahi, and A. Zilouchian, “Neural network model for a commericial PEM fuel cell system,” *Journal of Power Sources*, vol. 172, pp.749-759,2007.

[3] Y. Li, H. Wang, and Z. Dai “Using Artificial Neural Network to Control the Temperature of Fuel Cell,” *IEEE Conference*, 2006,pp. 2159-2162.

[4] A. P. E. M and M. G. Simoes, “Neural Optimal Control of PEMFC with Parametric CMAC Network,” *Industry Application Conferenc*e, 38th IAS Annual Meeting, vol. 2, pp.723-730, 2003.

[5] L. Zhang, M. Pan, S. Quan, Q. Chen, and Y. Shi, “Adaptive Neural Control Based on PEMFC Hybrid Model,” in *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian China, 2006.

[6] L. J. A. Dicks, “Fuel Cell Systems Explained,” *Wiley, London*, 2001.

[7] Z. Y. Dong, Z. Jianguo, G. Y. Guang, and J. Jianxun, “Control of Proton Exchange Membranes Fuel Cell Based on Fuzzy Logic,” in *proceedings of the 26th Chinese Control Conference*, July 26-31 200, Hunan China.

[8] R. F. Mann, J. C. Amphlett, and M. A. I. Hooper, et al, “Development and Application of a Generalized Steady-State Electrochemical Model for a PEM Fuel Cell,” *Journal of Power Sources*, vol. 86, pp.172-180, 2000.

[9] T. V. Nguen and R. E. White, “A Water and Heat Management Model for Proton-Exchange-membrane Fuel Cells,” *Journal of Electrochemical Society,* vol. 140, no. 8, pp. 2178-2186,1993.

[10] D. Ernst, et al, “Power System Stability Control: Reinforcement Learning Framework,” *IEEE Transaction on Power System February*, vol. 19, pp. 427-435, 2004.

[11] R. Bellman, “Dynamic Programming,” Princeton, NJ: Princeton Univ. Press, 1957S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” *IEEE Trans. on Neural Networks*, vol. 4, pp. 570-578, July 1993.

Mohammad Karimi was born in 1983 in Tehran, Iran. He received the B.Sc and M.S.c degrees both in electrical engineering from I.A.U, Ardebil and Ahar, Iran, in 2008 and 2011, respectively. He is working as the instructor in I.A.U. He has been involved with the FACTS devices (control and placement), distribution systems planning and optimization.

Mehdi Imanzadeh was born in 1981. He received the B.Sc and M.Sc. degrees both in Physical Chemistry from I.A.U, in 2004 and 2006, respectively. Mr. Imanzadeh is the Faculty member of Department of Applied Science -I.A.U, Parsabad Moghan Branch. His areas of interest in research are majorly in electrochemistry.

Payam Farhadi was born in 1985. He received the B.E.E and M.E.E. degrees both from I.A.U, in 2009 and 2011, respectively. Mr. Farhadi is an elite member of Young Researchers Club in Iran, also a member of the Faculty of Department of Electrical Engineering-I.A.U, Parsabad Moghan Branch. His areas of interest in research are distribution systems planning and optimization, distributed generation and FACTS Controllers. He has published more than 40 papers in International Journals and Conference Proceedings.

Noradin Ghadimi was born in 1985. He received the B.E.E and M.E.E. degrees both from I.A.U, in 2009 and 2011, respectively. His areas of interest in research are power system protection, modeling and analysis of distributed generations and renewable energies.