Applying reinforcement learning method to optimize an Energy Hub operation in the smart grid

Conference Paper · January 2015

1 author:

Aras Sheikhi
Sharif University of Technology

Applying Reinforcement Learning Method to Optimize an Energy Hub Operation in the Smart Grid

M. Rayati, A. Sheikhi, A. M. Ranjbar
Electrical Engineering Department
Sharif University of Technology
Tehran, Iran
m-Rayati@ee.sharif.edu

***Abstract*-- New days, the concepts of "Smart Grid" and "Energy Hub" have been introduced to improve the operation of the energy systems. This paper introduces a new conception entitling Smart Energy Hub (S. E. Hub), as a multi-carrier energy system in a smart grid environment. To show the application of this novel idea, we present a residential S. E. Hub which employs Reinforcement Learning (RL) method for finding a near optimal solution. The simulation results show that by applying the S. E. Hub model and then using the proposed method for a residential customer, running cost is reduced substantially. While, comparing with the classical ones, the RL method does not require any data about the environment and either equipment's parameters.**

***Index Terms*-- Smart Energy Hub (S. E. Hub), Smart Grids, Reinforcement Learning (RL), Energy Management System, Optimization**

I. INTRODUCTION

Common energy infrastructures such as electricity and natural gas systems are planned and operated independently. Motivated by different reasons, a number of recent publications suggest an integrated view of energy systems including multiple energy carriers; instead of focusing on single one [1]-[7]. In these publications, an exhaustive model describing multi-carrier energy systems properly is named Energy Hub [2]. The energy hub modeling and concept is proposed for the first time by Geidl. Within the hub, energy is converted using combined cooling, heating, and power (CCHP) systems, transformers, power-electronic devices, compressors, heat exchangers, and other equipment.

For controlling and optimizing the energy hub, different methods have been introduced and discussed [3]-[7]. Sheikhi et al. proposed a value based planning method for optimal sizing CCHP [3]. A cost-benefit analysis to determine the optimal size and operation of CCHP were given in [4] and [5]. Parisio et al. presented a robust optimization (RO) technique for solution of the energy hub optimization to satisfy the energy request [6]. Effect of demand response on the residential energy cost in a system equipped with micro-CCHP was investigated in [7].

From a number of recent publications [1]-[7], it can be seen that the daily loads are predetermined. Moreover, there are not any interlinks between customers and network. In other words, energy system is not implemented in the smart grid environment.

Smart grid is bidirectional flows of energy and communication that will enable an array of new functionalities and applications [8], [9]. In a smart grid environment, we have smart meters instead of conventional ones. Smart meters not only provide energy consumption readings but can also provide additional information on usage and have two-way communication capabilities. These features are two key developments that improve the effectiveness and capability of energy management systems (EMS) [10], [11].

In [12], a heuristic rules was used to manage the energy demand of a residential customer when plug-in hybrid electric vehicles (PHEVs) and local renewable energy resources are considered as a part of loads. Mohsenian-Rad in [13] proposed an automated optimization-based residential load control scheme by real time pricing (RTP) combined with inclining block rate (IBR). In [14], a network congestion game was used for controlling the power demands at peak hours. O'Neill in [15] scheduled domestic appliances usage by estimation of future energy prices and customer decisions impacts on long term costs. Bozchalui in [16] modeled the residential energy hub and solved this model to optimally control all major domestic energy loads with considering the customers preferences.

Detailed review of the technical literatures [1]-[16], obviously shows that most of the research papers only take into account particular appliances and their approaches are hardly applicable.

The current paper tries to deal with these weaknesses by proposing a new solution which is entitled smart energy hub (S. E. Hub). This solution allows homeowners to monitor and control energy use for reducing their electricity and natural gas bill. The proposed model is useful if it is solved effectively in a real-time frame in order to control all energy generation devices optimally and does not require any devices parameters. To achieve these purposes, the Reinforcement Learning (RL) method is applied.

The contents of this paper are organized into the following: The S. E. Hub concept and detailed specification of model is described in section II. In section III, we formulate the RL method and its application in optimizing S. E. Hub

978-1-4799-1785-3/15/$31.00 ©2015 IEEE

operation is described in section IV. Simulation results are discussed and compared with the conventional system in section V. Finally, the paper is concluded in section VI.

II. What is the "Smart Energy Hub"?

Today, demands increment is one of the greatest concerns in the energy systems with respect to the limited resources. Consequently, enhancing the efficiency of consumption is an inevitable goal, which leads us to a novel concept called "Smart Energy Hub" (S. E. Hub).

S. E. Hub uses one-way flows of natural gas and two-way flows of electricity and information in order to create an optimized unit. It means that S. E. Hub purchases natural gas and electricity while sells generated or stored electricity to the grid in proper time. Figure 1 depicts the overall view of S. E. Hub. Also, the operation of the conventional system is shown in Figure 2.

The concept of S. E. Hub can be used for a residential costumer with various appliances, e.g. PV array, PHEV, washing machine, dryer, refrigerator, heater, water heater, and lightening system. These appliances can be controlled or delayed by the agents in the S. E. Hub. In Figure 3, the residential S. E. Hub is shown.

An energy management system (EMS) optimizes the operational cost of the S. E. Hub. The EMS controls the appliances of the S. E. Hub with home area network (HAN) [17], [18].

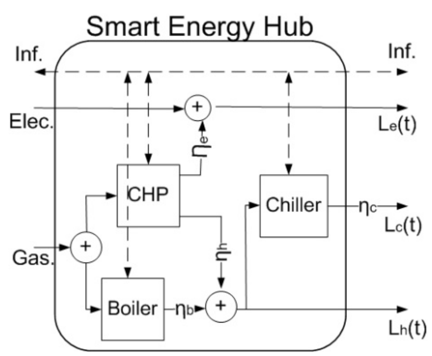

Figure 1. Overall view of a S. E. Hub

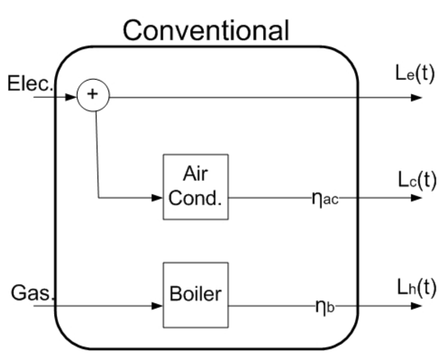

Figure 2. Overall view of a conventional system

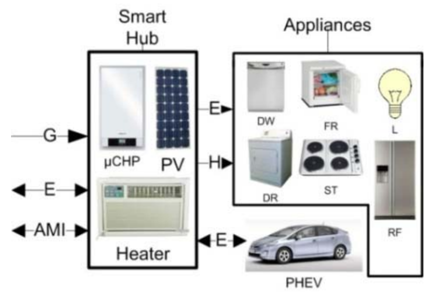

Figure 3. The residential S.E.Hub

III. Reinforcement Learning

For solving real-world problem, Reinforcement Learning (RL) method is suitable. RL is the study of programs that improve their performance and adapt an agent to an unknown environment by receiving rewards and punishments [19].

The RL agent interaction is characterized by the state ($s(t) \in S$), action ($a(t) \in A(s)$), and the reward ($r(t)$) signals. Here, S is the set of all the possible states; and $A(s)$ is the set of all possible actions in each state [20]. The overview of RL system is illustrated in Figure 4.

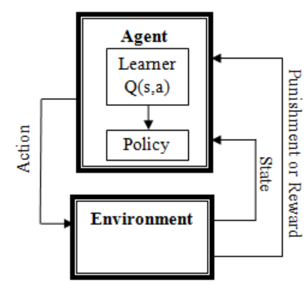

Figure 4. Overview of RL system

As shown in Figure 4, three main sub-elements of RL system can be identified as policy, reward or punishment, and Q-function [21].

Policy means which action is accomplished at each state. Here, $a^*(s)$ is the best action at state s.

$$\pi\ (s, a) = \begin{cases} 1 & a = a^*(s) \\ 0 & a \neq a^*(s) \end{cases} \tag{1}$$

Reward or punishment is something incentivizing customers to modify their policy in a RL problem.

Q-function is the value of taking an action in a state under policy π. It is denoted as $Q^\pi(s, a)$ which can be defined formally as (2).

$$Q^\pi(s, a) = E_\pi\{R(t)|\ s(t) = s, a(t) = a\} \tag{2}$$

where, $R(t)$ is the total reward that agent receives in the long run. Maximizing $R(t)$ or equivalently minimizing total punishments is the main objective of the agent. Total reward is described here as (3).

$$R(t) = \sum_{n=t}^{\infty} \gamma^{n-t} \times r(n) \tag{3}$$

in which, γ is a parameter in [0,1], calling discount rate.

The optimal action-value-function and optimal policy are denoted as $Q^*(s, a)$, π^* respectively [22]. To find the optimal policy, the Monte-Carlo method has been applied. The Monte-

Carlo has two main steps for finding optimal policy: evaluation and improvement. The overview of Monte-Carlo method is presented in Figure 5.

In evaluation step, the Q-function is filled with on-line interaction by the following formula.

$$Q(s,a) \cong \frac{\sum_{i=1}^{k} R_t}{k} \tag{4}$$

The improvement step is to find the best action in each state (i.e. policy) with (5).

$$a^*(s) = arg\max_a Q(s,a) \tag{5}$$

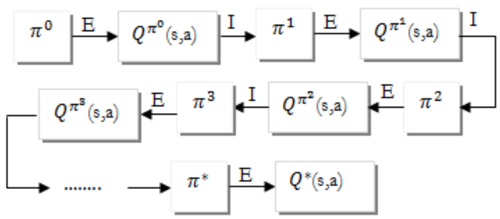

Figure 5. The Overview of Monte-Carlo method

Consequently, the optimization problem is solved by the Monte-Carlo method. However, if deterministic policy is used, there is no assurance regarding all actions selected to explore the optimal policy. Hence, a trade-off between exploitation and exploration is encountered [23]. To solve this new problem, the ε-soft on-policy method is used which updates action on the basis of the experience gained from executing policy [23]. This method is described further as:

$$\pi(s,a) = \begin{cases} 1 + \dfrac{\varepsilon}{|A(s)|} - \varepsilon & if \ \ a = a^* \\ \dfrac{\varepsilon}{|A(s)|} & if \ \ a \neq a^* \end{cases} \tag{6}$$

where, $|A(s)|$ represents the number of actions and ε is a parameter in $[0,1]$, calling probability of exploration.

IV. Method Descreption

To optimize the total cost of energy consumption in buildings, the EMS has been extensively applied. EMS can be programmed by various algorithms, for instance linear programming, which is described in following.

$$\min_{P_{gC}(t), P_{gB}(t)} \sum_t J(t) = \sum_t J_e(t) + J_g(t) \tag{7-a}$$

where,

$$J_e(t) = \left(L_e(t) - P_{gC}(t) \times \eta_e^C\right) \times Pr_e(t) \tag{7-b}$$

$$J_g(t) = \left(P_{gC}(t) + P_{gB}(t)\right) \times Pr_g \tag{7-c}$$

$subject\ to:\ \forall t$

$$P_{gC}(t) \times \eta_h^C + P_{gB}(t) \times \eta_h^B \geq L_h(t) + L_c(t)/\eta_C \tag{8-a}$$

$$P_{gC}(t) \leq \overline{P_{gC}} \tag{8-b}$$

$$P_{gB}(t) \leq \overline{P_{gB}} \tag{8-c}$$

here, $L_e(t)$ is the electrical load (kWh); $Pr_e(t)$ and Pr_g are the electricity and natural gas prices respectively ($/kWh); $P_{gC}(t)$ and $P_{gB}(t)$ are the natural gas inputting in the CHP and auxiliary boiler (kWh); $L_h(t)$ and $L_c(t)$ are the thermal and cooling loads (kWh); η_e^C and η_h^C are the electricity and thermal efficiency of CHP; η_h^B is the efficiency of auxiliary boiler; η_C is the efficiency of absorption chiller; and $\overline{P_{gC}}$, $\overline{P_{gB}}$ are the maximum input of CHP and auxiliary boiler.

In this paper, the actions of EMS for optimizing costs may affect not only the immediate reward but also all the subsequent rewards. To achieve this goal, learning algorithms are applied determining which action yields the highest reward by trial and error. These two characteristics –trial & error search and delayed reward– are the most important distinguishing features of RL algorithms [20].

Here, environment includes S. E. Hubs' elements, their parameters and customer's behaviors. Agent is the EMS device. Also, reward or punishment, state, and action at each time are as follow:

$$r(t) = -\left(P_{gC}(t) + P_{gB}(t)\right) \times Pr_g + P_{gC}(t) \times \eta_e^C \times Pr_e(t) \tag{9-a}$$

$$s(t) = [Pr_e(t)] \tag{9-b}$$

$$a(t) = [P_{gC}(t)] \tag{9-c}$$

The RL algorithm is works properly when the state representation has the Markov property. The state has Markov property, if it succeeds in retaining all relevant information. For example, a checkers position (the current configuration of all the pieces on the board) would serve as a Markov state, because it summarizes everything important about the complete sequence of positions that led to it [20].

In this problem, the behavior of the customer and the electricity price is periodic such as diurnally or weekly cycles; hence, this algorithm can be employed [15].

V. Simulation

In this section, simulation results and performance assessment of the proposed algorithm have been illustrated. Here, the classical model was employed for simulating stochastic load profiles [24]. The functions $m(t_i)$ and $s(t_i)$ are required that indicate mean and variance of load at time t_i, respectively. Hence, $y(t_i)$ or electricity consumption between time t_{i-1} and t_i is as (10).

$$y(t_i) = m(t_i) + \sigma_i \tag{10}$$

where, the index $i = 1, .., n$ denotes the data frequency; σ_i is independent and identically distributed error terms with zero mean and $s(t_i)$ variance; $t_i = i/n$ is a time index; and n is the total number of observations.

We consider a residential S. E. Hub in a time-of-use (TOU) pricing market for electricity and fixed-rate-plan (FRP) for natural gas [16]. Electricity and natural gas energy prices are shown in Figure 6.

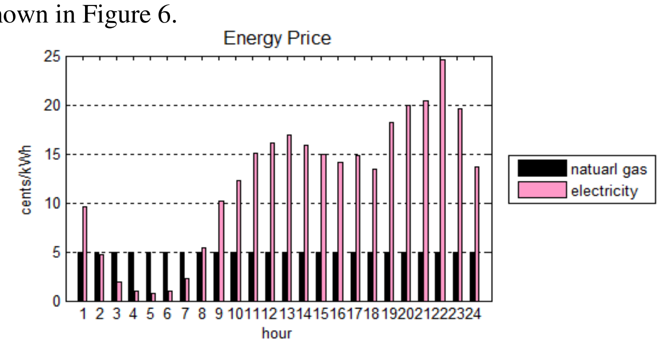

Figure 6. Electricity and natural gas prices

Electrical, thermal, and cooling loads for a sample day in summer and winter are depicted in Figure 7.

CHP, auxiliary boiler, absorption chiller, and transformer specifications are summarized in table I.

S. E. Hub outputs' power, are shown in Figure 8. In table II, operations of the conventional system and S. E. Hub have been compared with each other.

Q-function of EMS learner is shown in Figure 9. In each step, it determines which action should take place on specific state (i.e. policy). For instance, Figure 9 indicates that CHP would generate higher in prohibitive electricity price conditions and would decline its generation in low price electricity. The parameters of proposed method are depicted in Table III.

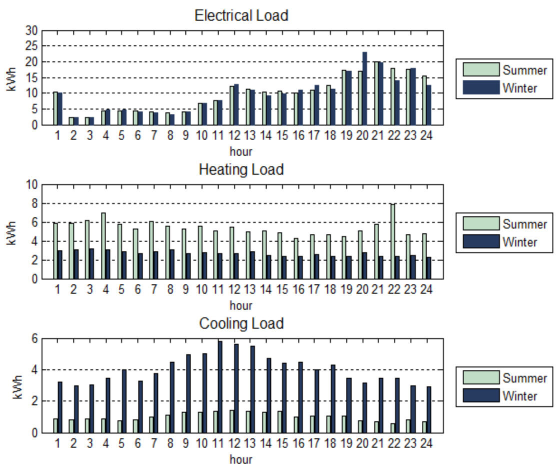

Figure 7. S. E. Hub loads.

TABLE I. DEVICES SPECIFICATIONS

Auxiliary Boiler	***Efficiency (%)***	96
	Size (kW)	10
Absorption Chiller	***Efficiency (%)***	65
	Size (kW)	5
Transformer	***Efficiency (%)***	97
	Size (kVA)	20
CHP	***Electrical eff. (%)***	37
	Thermal eff. (%)	43
	Size (kW)	10

TABLE II. SUMMARIZED INFORMATION OF COMPARISON BETWEEN THE SMART ENERGY HUB (S.E.HUB) AND CONVENTIONAL SYSTEM (CONV.)

		Winter	**Summer**
Consumption (KWh)	***Conv.***	234.9	246.7
	S.E.Hub	242.8	238.6
Natural Gas (kWh)	***Conv.***	144.2	72.1
	S.E.Hub	375.5	470.9
Generation (KWh)	***Conv.***	0	0
	S.E.Hub	114.71	152
Energy bill ($)	***Conv.***	47.6	58.0
	S.E.Hub	39.2	38.7

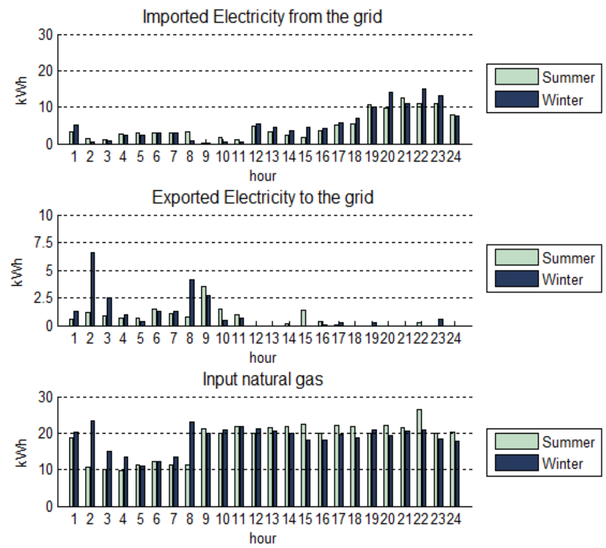

Figure 8. S. E. Hub Input

TABLE III. PROPOSED METHOD PARAMETERS

Discount rate	γ	0.05
Exploration rate	ε	0.1

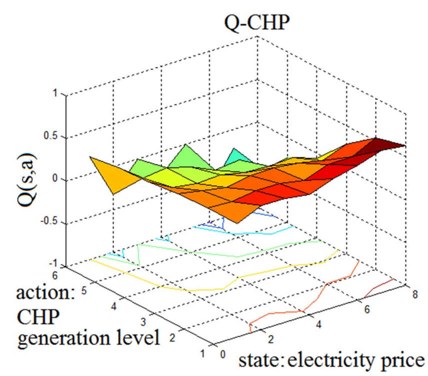

Figure 9. Q-function of CHP

As shown in Figures 7, 8 and table II, by applying this approach, the peak load is reduced from 20.3kWh to 13.5kWh (i.e. 33% less). Also, the total energy bill reduces from 52.8$ to 38.95$ (i.e. 26% improvement).

VI. CONCLUSION

This paper introduces "smart energy hub" (S.E.Hub) as a new concept and proposes reinforcement learning (RL) algorithm as a practical method to find near optimal electricity and natural gas consumptions of this system. This algorithm does not require any detail information about dynamic of the system, and converges in a very short time. The simulation results confirm the efficiency of the model and algorithm and demonstrate up to 30% reduction in the electricity and gas bill.

ACKNWLEDGEMENT

The authors would like to express their deepest gratitude to the Iran's National Elite Foundation, for the financial support during this study.

REFERENCES

[1] J. Harrison, “Micro combined heat & power (CHP) for housing”, in *proc. of Int’l Conf. on Sustainable Energy Tech.*, Nottingham, UK, June 2004.

[2] M. Geidl, G. Koeppel, P. Favre-Perrod, B. Klöckl, G. Andersson, and K. Fröhlich, “Energy Hubs for the futures”, *IEEE Power & Energy Mag.*, vol. 5, no. 1, pp.24–30, Jan./Feb. 2007.

[3] A. Sheikhi, A. Ranjbar, and H. Oraee, “Financial analysis and optimal size and operation for a multicarrier energy system”, *Energy and buildings*, vol. 48, pp. 71–78, Jan. 2012.

[4] A. Sheikhi, A. Ranjbar, and F. Safe, “A novel method to determine the best size of CHP for an energy hub system”, in *proc. of Int’l Conf. on renewable energy and power quality*, Sharjah, UAE, Nov. 2011.

[5] A. Sheikhi, A. Ranjbar, F. Safe, and M. Mahmoody, “CHP optimized selection methodology for an energy hub system”, in *proc. of 10th Int’l Conf. on Environment and Electrical Engineering*, Rome, Italy, May 2011.

[6] A. Parisio, C. Del Vecchio, and G. Velotto, “Robust Optimization of operations in energy hub”, in *proc. of 50th IEEE Conf. on Decision and Control and European Control Conf.*, Orlando, USA, Dec. 2011.

[7] M. Houwing, R. Negenborn, and B. DeSchtter, “Demand Response with Micro-CHP Systems”, in *proc. of IEEE*, vol. 99, no. 1, pp. 200–213, Jan. 2011.

[8] National Institute of Standards and Technology, NIST framework and roadmap for smart grid interoperability standards, release1.0, pp. 20–100, Available at http://www.nist.gov/public affairs/releases/upload/sma-rtgrid interoperability final.pdf., Jan. 2010.

[9] Xi Fang, S. Misra, G. Xue, and D. Yang, “Smart Grid – The New and Improved Power Grid: *A Survey”, IEEE Communications surveys & tutorials*, vol. 14, no. 4, pp. 944–980, Oct. 2012.

[10] Y. Guo, M. Pan, and Y. Fang, “Optimal power management of residential customers in the smart grid”, *IEEE Trans. on parallel and distribute systems*, vol. 23, no. 9, pp. 1593–1606, Sept. 2012.

[11] A. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, “Autonomous demand-side management based on game theoretic energy consumption scheduling for the future smart grid”, *IEEE Trans. on Smart Grid*, vol. 1, no. 3, pp. 320–331, Nov. 2010.

[12] Oviedo R.J.M. , Zhong Fan, Gormus, S., Kulkarni, P., and Kaleshi, D., “Residential energy demand management in smart grids”, in *proc. of Transmission and Distribution Conf. and Exposition*, Orlando, USA, May 2012.

[13] A.H. Mohsenian-Rad and A. Leon-Garcia, “Optimal residential load control with price prediction in real-time electricity pricing environments”, *IEEE Trans. on Smart Grid*, vol. 1, no. 2, pp. 120–133, Sept. 2011.

[14] C. Ibars, M. Navarro, and L. Giupponi, “Distributed demand management in smart grid with a congestion game”, in *proc. of IEEE SmartGridComm*, Gaithersburg, USA, Oct. 2010.

[15] D. O’Neill, M. Levorato, A. Goldsmith, and U. Mitra, “Residential demand response using reinforcement learning”, in *proc. of IEEE SmartGridComm*, Gaithersburg, USA, Oct. 2010.

[16] M. Chehreghani Bozchalui, A. Hashmi, H. Hassen, C.A. Cañizares, and K. Bhattacharya, “Optimal operation of residential energy hubs in smart grids”, *IEEE Trans. on smart grid*, vol. 3, no. 4, pp. 1755–1766, Oct. 2012.

[17] G. Song, F. Ding, W. Zhang, and A. Song, “A wireless power outlet system for smart home”, *IEEE Trans. on Consumer Electronics*, vol. 54, no. 4, pp.1688–1691, Nov. 2008.

[18] Dae-Man Han and Jae-Hyun Lin, “Smart home energy management system using IEEE 802.15.4 and ZigBee”, *IEEE Trans. on Consumer Electronics*, vol. 56, no. 3, pp. 1403–1410, Oct. 2010.

[19] K. Miyazaki, M. Yamamura, S. Kobayashi, “K-Certainty Exploration Method: an act .ion selector to identify the environment in reinforcement learning”, *Artificial Intelligence*, vol. 91, pp. 155–171, 1997.

[20] R.S. Sutton, “Introduction: the challenge of reinforcement learning”, in *Machine Learning, Kluwer academic publishers*, Netherlands, pp. 225–227, 1992.

[21] L. Kaelbling, M.L. Littman, and A.W. Moore, “Reinforcement learning: a survey”, *journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[22] S. Singh, T. Jaakkola, M.L. Littman, and C. Szepesvary, “Convergence results for single step on-policy reinforcement learning algorithms”, in *Machine Learning, Kluwer academic publishers*, Netherlands, pp. 287–308, 2000.

[23] R.S.Sutton, A.G.Barto, Reinforcement learning: an introduction, *MIT press*, London, England, 2005.

[24] M.Manera, A.Marzullo, “Modeling the load curve of aggregate electricity consumption using principal components”, *Environmental Modeling & Software*, vol. 20, pp. 1389-1400, 2005.