Taylor & Francis
Taylor & Francis Group

International Journal of Parallel, Emergent and Distributed Systems

ISSN: 1744-5760 (Print) 1744-5779 (Online) Journal homepage: http://www.tandfonline.com/loi/gpaa20

Optimising operational cost of a smart energy hub, the reinforcement learning approach

Mohammad Rayati, Aras Sheikhi & Ali Mohammad Ranjbar

International Journal of Parallel, Emergent and Distributed Systems, 2015
Vol. 30, No. 4, 325–341, http://dx.doi.org/10.1080/17445760.2014.974600

Optimising operational cost of a smart energy hub, the reinforcement learning approach

Mohammad Rayati, Aras Sheikhi* and Ali Mohammad Ranjbar

Sharif University of Technology, Tehran, Iran

(*Received 25 June 2013; Revised 24 June 2014; final version received 6 October 2014*)

The concept of smart grid (SG) has been introduced to improve the operation of the power systems. In modern structures of power systems, different reasons prompt researchers to suggest integrated analysis of multi-carrier energy systems. Considering synergy effects of the couplings between different energy carriers and utilising intelligent technologies for monitoring and controlling of energy flow may change energy system management in the future. In this paper, we propose a new solution which is entitled 'smart energy hub' (SEH) that models a multi-carrier energy system in a SG. SEH solutions allow homeowners to manage their energy consumption to reduce their electricity and gas bill. We present this concept for a residential customer by an 'energy management system' which uses reinforcement learning algorithm and Monte Carlo estimation method for finding a near optimal solution. The simulation results show that by using this concept and then applying the algorithm for a residential customer, running costs are reduced up to 40% while keeping the household owner's desired comfort levels.

Keywords: smart energy hub; smart grids; reinforcement learning; energy management system; optimisation

1. Introduction

In the past, common energy infrastructures such as electricity and natural gas systems were 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 a single energy carrier.[1–7] Coupling of these infrastructures is set up by micro combined heat and power (μCHP) which converts natural gas to electricity and heating.[1] Approaching from optimising gas and electricity networks distinctly to a whole system regards both networks' results as an optimal solution instead of two sub-optimal solutions.[2]

More recently, the combined modelling and analysis of energy systems including multiple energy carriers has been addressed in a number of publications, e.g. [1–7]. In these publications, an exhaustive model describing multi-carrier energy systems properly was named energy hub.[2,3] The 'energy hub' modelling and concept was proposed for the first time by Geidl. Within the hub, energy is converted and conditioned using CHP technology, transformers, power-electronic devices, compressors, heat exchangers and other equipment. Sheikhi et al. [3] proposed a value-based planning method for optimal sizing combined cooling, heating and power (CCHP) placement based on the energy hub concept. Refs [4,5] give a cost–benefit analysis to determine the optimal size and operation of CHP and CCHP. In [6], a robust optimisation technique is presented for solution of energy hub optimisation to satisfy the energy request while

*Corresponding author. Email: sheikhi@ee.sharif.edu

minimising a cost function. Ref. [7] investigated the effect of demand response on the residential energy cost in a system equipped with μCHP.

From a number of recent publications,[1–7] it can be seen that daily loads are predetermined and there are no interlinks between users and network. In other words they did not implement energy system in the smart grid (SG) environment.

SG is bidirectional flows of energy and two-way communication that will enable an array of new functionalities and applications.[8] Due to various advantages such as energy efficiency, reliability and self-monitoring, the SG technology has been emerging as the next-generation intelligent power grid system.[9] In a SG environment, we have ‘smart meters’ instead of conventional meters. 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 (EMSs).[10]

From the perspective of residential customers, the aim of smart management system is how to minimise electricity pricing bill,[11] and from the view of network manager, the goal is how to minimise the cost of generating energy by the sources.[12]

The vulnerability of power grid systems to malicious attacks is one of the most pressing problems faced concerning power grid systems.[13] The natures of privacy leakages in demand response programme are investigated and the potential privacy threat models are explored in [14].

Ref. [15] used heuristic rules to manage the energy demand in a residential consumer when plug-in hybrid electric vehicles (PHEVs) and local renewable energy resources are considered as a part of loads. In [16], the blocking probability of the communication-based PHEV load management scheme is analysed, and the optimal required capacity to supply all of the PHEVs’ loads by considering grid limitations has been calculated. Mohsenian-Rad and Leon-Garcia [17] proposed an automated optimisation-based residential load control scheme by real-time pricing combined with inclining block rate. Ref. [18] used network congestion game to control the power demand at peak hours. O’Neill et al. [19] scheduled domestic devices usage by estimation of impact of future energy prices and of consumer decisions on long-term costs. Bozchalui et al. [20] modelled the residential energy hub and solved this model to optimally control all major domestic energy loads with considering the consumers preferences.

Detailed review of the technical literatures [1–20] obviously shows that most of the research papers only take into account particular appliances and are designed for one particular objective (e.g. peak demand reduction or energy consumption minimisation). In those papers user’s preferences and satisfaction level have not properly been designed.

This paper tries to deal with these weaknesses by proposing a new solution which is entitled ‘smart energy hub’ (SEH) that models SG in multi-carrier energy systems. SEH solutions allow homeowners to monitor and control energy use in real time and therefore manage their energy consumption to reduce their electricity and gas bill. The proposed model is useful if it is solved effectively in a real-time frame in order to control all major energy loads, storage and generation devices optimally by considering the customers comfort level in the same time. To achieve these purposes, the reinforcement learning (RL) method is applied.

The contents of this paper are organised into the following five sections: the SEH concept and detailed specification of model are described in Section 2. In Section 3, we formulate the RL method and its application in optimising SEH operation. Simulation results are discussed and compared with the conventional system in Section 4. Finally, conclusion is drawn in Section 5.

2. What is the 'SEH'?

The term 'SEH' is used as a unit in a smart energy infrastructure, where multiple energy carriers, e.g. natural gas and electricity, can be converted, conditioned and stored. SEH is an energy hub, which includes various appliances, loads and energy production systems that uses a smart meter and two-way communication links for the optimal operation in order to be smarter.

Today, one of the great concerns in the energy field is the more and more increasing demand of energy with respect to its limited resources. Consequently, increasing the efficiency of energy production systems is an inevitable approach.

CHP generates electricity and simultaneously reuses wasted heat in order to condition room temperature and warm up running water. It is forecasted that 54.5% of total energy of a house will be consumed for space and water heating in 2035.[25] So we consider it as the heart of the SEH in our model.

A SEH uses one-way flows of natural gas and two-way flows of electricity and information in order to create an optimised unit. It means SEH purchases gas and electricity and sells generated or stored electricity to the grid in proper time. Information has a bidirectional flow and the SEH receives information from the grid to optimise its behaviour and shares its information with the grid to promote grid behaviour as well. Figure 1 depicts the overall view of a SEH.

The residential SEH includes various appliances [e.g. washing machine (WM), dryer (DR), refrigerator (RF), heating, water heating, lightening, photo voltaic (PV) array, PHEV, CHP and other electronic appliances]. Each appliance has different operational constraints, which are presented by Bozchalui et al. [20]. Recently, Zigbee and IEEE 802.15.4 are used for energy management in homes and commercial building automation.[24]

An EMS optimises the operational cost of the SEH. In this paper, the presented method is based on the machine learning concept. In Figure 2, the residential SEH is shown.

The EMS controls the appliances of the SEH with home area network.[24] Ref. [22] presents the design and implementation of ZigBee based on wireless controllable power outlet architecture to build home automation networks.

The proposed mechanism of the EMS is portrayed in Figure 3. Number of manipulated variables, which is shown in following figure, is constant and dependent on degree of freedom.[23] Because of some restrains, some of the manipulated variables are not free to control by EMS all the time. Hence to resolve the explained issue, a selector has been applied to determine whether each EMS outputs commands to the corresponding

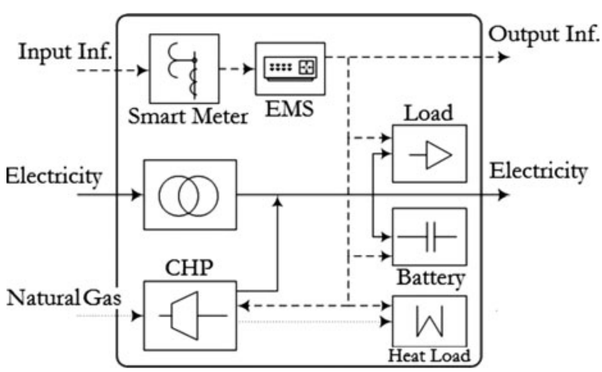

Figure 1. Overall view of SEH.

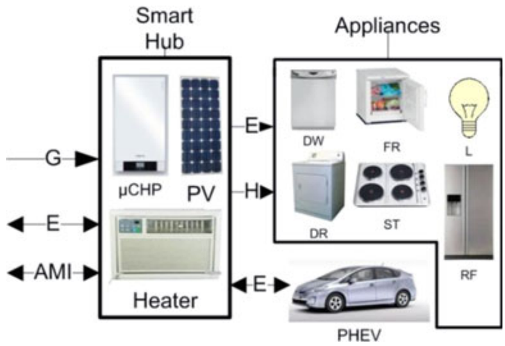

Figure 2. The residential SEH.

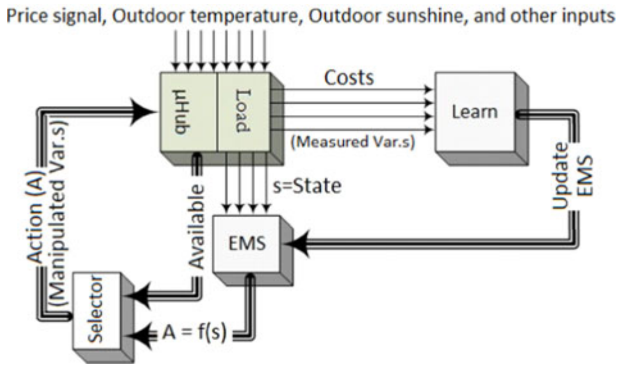

Figure 3. The proposed mechanism of the EMS.

manipulated variable or not. For instance, if a load such as WM starts its operation, it requires continuing its activity for some minutes. So a selector bypasses EMS output for this device in this duration.

Price signal, outdoor temperature, outdoor luminance, activity level and user preferences are the most important input variables which SEH functions are based on.

3. Method description

Consider a residential SEH in a time-of-use pricing programme for electricity and fixed rate plan for natural gas.[20] Suppose a home equipped with several electrical devices such as WM, DR, dishwasher (DW), stove (ST), heater (HT), air conditioner (AC), water heater (WH), RF, freezer (FR) and other electronic appliances.

The total energy consumption of all appliances can be calculated as follows (1):

$$\begin{aligned} l(t) = {} & l_{\mathrm{WM}} z_{\mathrm{WM}}(t) + l_{\mathrm{DR}} z_{\mathrm{DR}}(t) + l_{\mathrm{DW}} z_{\mathrm{DW}}(t) \\ & + l_{\mathrm{ST}} z_{\mathrm{ST}}(t) + l_{\mathrm{HT}} z_{\mathrm{HT}}(t) + l_{\mathrm{AC}} z_{\mathrm{AC}}(t) + l_{\mathrm{WH}} z_{\mathrm{WH}}(t) \\ & + l_{\mathrm{RF}} z_{\mathrm{RF}}(t) + l_{\mathrm{FR}} z_{\mathrm{FR}}(t) + \frac{\mathrm{lux}_{\mathrm{HL}}(t)}{\eta_{\mathrm{HL}}} + \frac{\mathrm{lux}_{\mathrm{RM}}(t)}{\eta_{\mathrm{RM}}} + \frac{\mathrm{lux}_{\mathrm{KT}}(t)}{\eta_{\mathrm{KT}}} + l_{\mathrm{UCL}}(t), \end{aligned} \tag{1}$$

where $l(t)$ is total load at time t; $z_i(t)$ is binary variable denoting states of device i at time t: ON, OFF; l_i is energy consumption of device i, when appliance i is ON; $\mathrm{lux}_z(t)$ is light flux for zone z (HL, hall; RM, room; KT, kitchen) at time t in lumens and η_z (in lm/W) can be expressed as the ratio of irradiated light flux (in lumens) and the received light active power in W. In [25], the efficacies of different light sources have been described. In this study, we suppose the efficacy of warm white LED as 80 lm/W. $l_{\mathrm{UCL}}(t)$ is firm loads such as TV, computer, DVD, etc. These appliances must work without any interruption at time t.

Furthermore, the SEH has a number of electricity generation devices such as PV, CHP and PHEV which generate heat and electricity as given in (2) and (3):

$$g(t) = g_{\mathrm{PV}}(t) + g_{\mathrm{CHP}}(t) + g_{\mathrm{PHEV}}(t), \tag{2}$$

where

$$g_{\mathrm{CHP}}(t) = g_{\mathrm{CHP-h}}(t) + g_{\mathrm{CHP-e}}(t), \tag{3}$$

where $g(t)$ is the total electricity generation at time t, $g_i(t)$ is the electricity generation of device i at time t and $g_{\mathrm{CHP-e}}(t)$ and $g_{\mathrm{CHP-h}}(t)$ are the generated heat and electricity of CHP. Electrical efficiency of CHP ($\eta_{\mathrm{ge}}^{\mathrm{CHP}}$) is about 15% and its heating efficiency ($\eta_{\mathrm{gh}}^{\mathrm{CHP}}$) is considered to be about 80%.[5]

Today to optimise the total cost of buildings, the EMS has been extensively applied. In this paper, this method will be implemented. Costs are divided into four parts, which is described in (4)–(7) and (11):

$$\mathrm{Min}\, J = J_1 + J_2 + J_3 + J_4, \tag{4}$$

$$J_1 = \text{electricity cost} = \sum_t (l(t) + l_{\mathrm{PHEV}}(t) - g(t)) \times P_{\mathrm{r}}(t), \tag{5}$$

$$J_2 = \text{natural gas cost} = \sum_t lg_{\mathrm{CHP}}(t) \times P_{\mathrm{g}}, \tag{6}$$

$$J_3 = \text{dissatisfaction cost} = \sum_t y(t) \times P_{\mathrm{d}} + \mathrm{ch}'(t) \times P_{\mathrm{c}} \times \delta(t - 8\,\mathrm{am}), \tag{7}$$

where $P_{\mathrm{r}}(t)$ is the electricity price at time t, $lg_{\mathrm{CHP}}(t)$ is the CHP natural gas input at time t in m^3, P_{g} is natural gas price in \$/m^3, $y(t)$ is dissatisfaction function at time t, $\mathrm{ch}'(t)$ is PHEV remaining energy level at time t, P_{d} is dissatisfaction price relating to delay loads and P_{c} is dissatisfaction price relating to PHEV.

Dissatisfaction cost consists of two terms. The first term is related to the consumers' resent cause of appliances delay and the second one models PHEV storage level, which affects dissatisfaction. Consumers will prefer that devices complete their duty as soon as possible; thus making any delay in running reduces their satisfaction. To model consumers' behaviour dissatisfaction function, one can define it as (8) and (9) [19]:

$$y(t+1) = \theta y(t) + (1 - \theta) x(t), \tag{8}$$

where θ is a parameter between 0 and 1 which weights up the past memories of dissatisfaction. $x(t)$ is the average of delays and is defined as

$$x(t) = \left[\sum \beta_i \frac{w_i(t)}{T_{\max,i}}\right] \Big/ \sum \beta_i, \tag{9}$$

where β_i is an integer parameter denoting importance of device i, $w_i(t)$ is the delay time of appliance i at time t and $T_{\max,i}$ is the maximum allowable of $w_i(t)$.

Furthermore, all consumers need to have a PHEV with a full battery for daily use. Thus if the residual charge level of PHEV's battery is not zero, the satisfaction level is decreased.

$$\mathrm{ch}'(t) = \frac{[\text{Max Level} - \mathrm{ch}(t)]}{\text{Max Level}}. \tag{10}$$

In the above equations, we denote the price of delay dissatisfaction by P_d and PHEV dissatisfaction by P_c.

$$J_4 = \mathrm{CO_2}\text{ emission cost} = \sum_t M_{\mathrm{CO_2}}(t) \times P_{\mathrm{CO_2}}, \tag{11}$$

where $P_{\mathrm{CO_2}}$ is the emission price and $M_{\mathrm{CO_2}}$ is the mass of CO_2 emission (in kg) which is shown below:

$$M_{\mathrm{CO_2}}(t) = (l(t) + l_{\mathrm{PHEV}}(t) - g(t)) \times \alpha_{\mathrm{grid}} + g_{\mathrm{CHP}}(t) \times (\alpha_{\mathrm{CHP}} - \alpha_{\mathrm{grid}}), \tag{12}$$

where α_{grid} is CO_2 emission rate of grid and α_{CHP} is CO_2 emission rate of CHP. In this article, the actions of EMS for optimising costs may affect not only the immediate reward but also the next situation as well as all the subsequent rewards. Therefore, learning algorithms are used, which are not told what action should take place; but instead, they must discover what actions yield the highest reward by their trial. These two characteristics – trial and error search and delayed reward – are the most important distinguishing features of RL algorithms.[26–29]

The algorithm converges if and only if the problem is modelled as Markov decision process (MDP) and has Markov chain conditions. A MDP consists of a set of states S; a set of actions A; a state transition function $T(s, a, s')$, which is equal to probability of making a transition from state s to state s' using action a.

The model has Markov chain if the state transitions are independent of any previous environment state or agent actions.[26] In this problem, the behaviour of the customer is periodic such as diurnally or weekly cycles; thus this algorithm can be used.[19] The transition probabilities associated with this Markov chain can vary with different residence and are thus assumed to be unknown and are just known by learners.[26] The overview of a MDP is illustrated in Figure 4.

To implement this algorithm, due to optimisation costs of a SEH, four independent controllers are used, which is, respectively, depicted and described in Figure 5.

3.1 Delaying loads controller

Delaying loads controller (DLC) is employed to control WM, DR, DW, ST, lightening, AC, HT, WH, RF and FR. The reward signal is given as

$$r_{\mathrm{DLC}}(t) = -l(t) \times P_\mathrm{r}(t) - y(t) \times P_\mathrm{d}. \tag{13}$$

The reward signal reveals us the fact if the price of electricity or dissatisfaction increases, this controller is punished by negative reward; thus the states and actions should

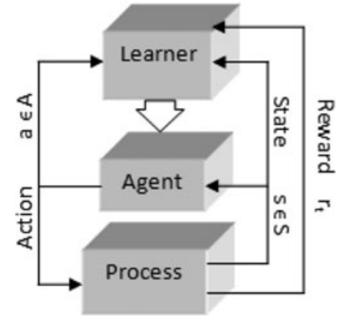

Figure 4. Overview of a MDP.

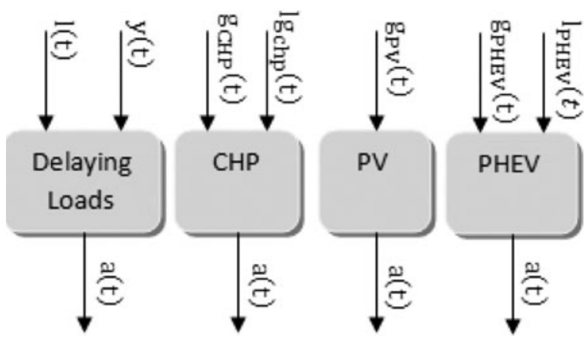

Figure 5. Applied controllers for optimising costs of SEH.

be given as

$$S_{\mathrm{DLC}}(t) = [P_{\mathrm{r}}(t), y(t)], \tag{14}$$

$$A_{\mathrm{DLC}}(t) = \{a_i(t) | a_i(t) \in [-1, 0, 1]\}, \tag{15}$$

where if $a_i(t) = 1$, the EMSs delay device i and if $a_i(t) = -1$, device i keeps running its remaining task.

3.2 CHP controller

CHP controller (CHC) is employed to control generation level of CHP. The CHP should generate in high level if electric price is prohibitive and vice versa. Therefore, the reward, states and actions should be

$$r_{\mathrm{CHC}} = -lg_{\mathrm{CHP}}(t) \times P_{\mathrm{g}} + g_{\mathrm{CHP}}(t) \times P_{\mathrm{r}}(t), \tag{16}$$

$$S_{\mathrm{CHC}}(t) = [P_{\mathrm{r}}(t)], \tag{17}$$

$$A_{\mathrm{CHC}}(t) = \{a_{\mathrm{CHP}}(t)\} = g_{\mathrm{CHP}}(t). \tag{18}$$

3.3 PV controller

The PV controller (PVC) must benefit from sunlight energy in the maximum margin. Thus it should generate in the appropriate condition. To accomplish this objective, the reward, states and actions shall be given as

$$r_{\mathrm{PVC}} = g_{\mathrm{PV}}(t) \times P_{\mathrm{r}}(t), \tag{19}$$

$$S_{\mathrm{PVC}}(t) = [P_{\mathrm{r}}(t),\ \mathrm{lux}(t)], \tag{20}$$

$$A_{\mathrm{PVC}}(t) = \{a_{\mathrm{PV}}(t)\} = g_{\mathrm{PV}}(t). \tag{21}$$

3.4 PHEV controller

To simplify the process, it is assumed that the PHEV is plugged out at 8 am with approximately full battery and is plugged in quite empty at 4 pm; in this interval, the battery can store electric energy at low price hours and use its energy at high price hours. To achieve this scenario, the reward, states and actions must be:

$$r_{\mathrm{PHC}}(t) = -l_{\mathrm{PHEV}}(t) \times P_{\mathrm{r}}(t) + g_{\mathrm{PHEV}}(t) \times P_{\mathrm{r}}(t) - \mathrm{ch}'(\mathrm{t}) \times P_{\mathrm{c}} \times \delta(t - 8\,\mathrm{am}), \tag{22}$$

$$S_{\mathrm{PHC}}(t) = [P_{\mathrm{r}}(t),\ \mathrm{ch}'(\mathrm{t}),\ t], \tag{23}$$

$$A_{\mathrm{PHC}}(t) = \{g_{\mathrm{PHEV}}(t),\ l_{\mathrm{PHEV}}(\mathrm{t})\}. \tag{24}$$

For each of the four controllers defined above, recall that a policy π is a mapping from each state, $s\epsilon S$ and action $a\epsilon A(s)$, to the probability $\pi(s, a)$ of taking action a when it is in state s.

For the MDP, the value of a state s under a policy π is denoted as $V^{\pi}(s)$, and the value of taking action a in state s under a policy π is denoted as $Q^{\pi}(s, a)$ which can be defined formally as

$$V^{\pi}(s) = E_{\pi}\{R_t | s_t = s\} = E_{\pi}\left\{\sum_{n=t}^{\infty} \gamma^{n-t} r_n | s_t = s\right\}, \tag{25}$$

$$Q^{\pi}(s, a) = E_{\pi}\{R_t | a_t = a,\ s_t = s\} = E_{\pi}\left\{\sum_{n=t}^{\infty} \gamma^{n-t} r_n | a_t = a,\ s_t = s\right\}, \tag{26}$$

where γ is a parameter, $0 \le \gamma \le 1$, called the discount rate. This parameter is varying for each controller.

Our goal is to find a policy that achieves numerous rewards over the long run. The optimal value function, the optimal action-value function and optimal policy are denoted as $V^*(s)$, $Q^*(s, a)$ and π^*.[28]

To find the optimal policy, the Monte Carlo method is used, which requires us only to experience sample sequences from online or interaction with an environment. The Monte Carlo has two steps for finding optimal policy: evaluation and improvement. The overview of Monte Carlo method is presented in Figure 6.

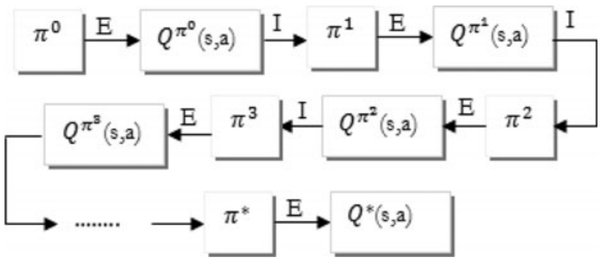

Figure 6. The overview of Monte Carlo method.

In evaluation step, Q function is filled with online interaction by the following formula:

$$Q(s,a) \cong \frac{\sum_{i=1}^{k} R_t}{k}. \tag{27}$$

The improvement step is to find policy π with the algorithm below.

$$\pi(s) = \arg\max_{a} Q(s,a). \tag{28}$$

Consequently, the problem of optimisation is solved by Monte Carlo method of Q-learning algorithm. However, if more attention is devoted, it is found clearly 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

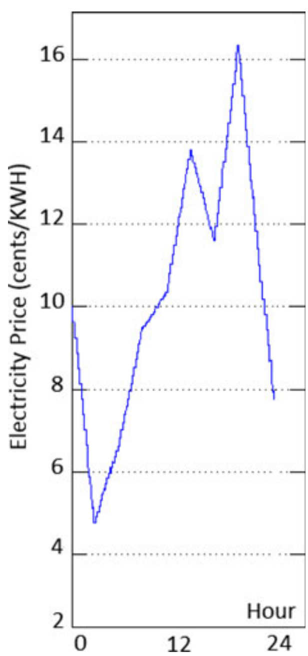

Figure 7. Electricity price.

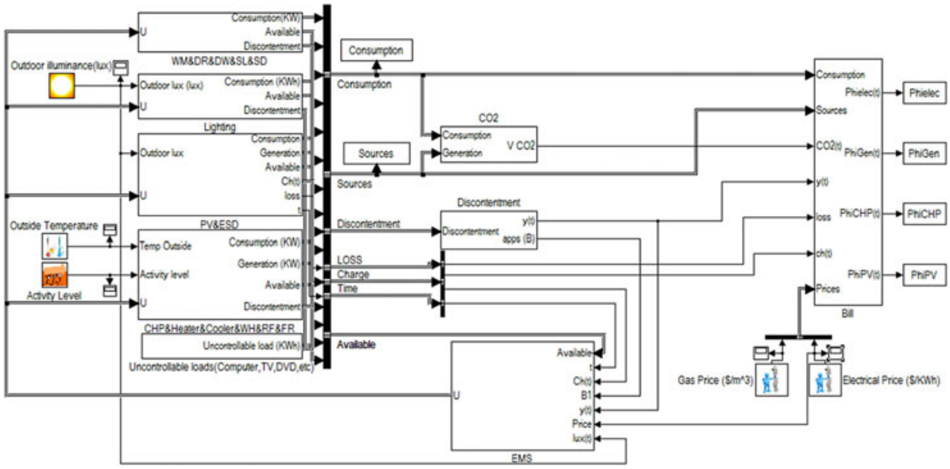

Figure 8. Simulated SEH in 'Matlab Simulink'.

encountered.[27] 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.[27] This method is described further as

$$\pi(s,a) = \begin{cases} 1 + \frac{\varepsilon}{|A(s)|} - \varepsilon, & \text{if } a = a^*, \\ \frac{\varepsilon}{|A(s)|}, & \text{if } a \neq a^*, \end{cases} \tag{29}$$

Table 1. Tehran climate average of January.

Average maximum temperature (F)	46
Average minimum temperature (F)	30
Mean (F)	38
Cloudy days	6
Wet days	5
Fair days	20
Average sky cover	0.3

Table 2. PV, PHEV and CHP specifications.

PV system	Array	No. of modules	3
		Rated power (W)	80 × 3
		Voltage (V)	17.86
		Price ($)	238.56 × 3
	Battery	Ch current (A)	8.5
		Capacity (Ah)	35
		Voltage (V)	12
		Price ($)	65.5
PHEV (20 miles)		Capacity (kW h)	6.7
		Power (W)	1251
		Consumption (W h/mile)	300
CHP [35]		*E* efficiency (%)	15
		H efficiency (%)	80
		E output (kW)	1

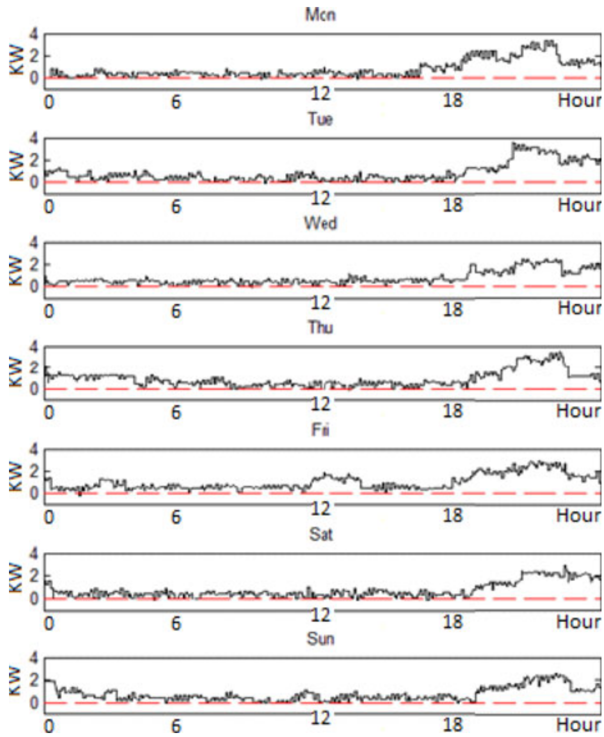

Figure 9. SEH output before applying EMS.

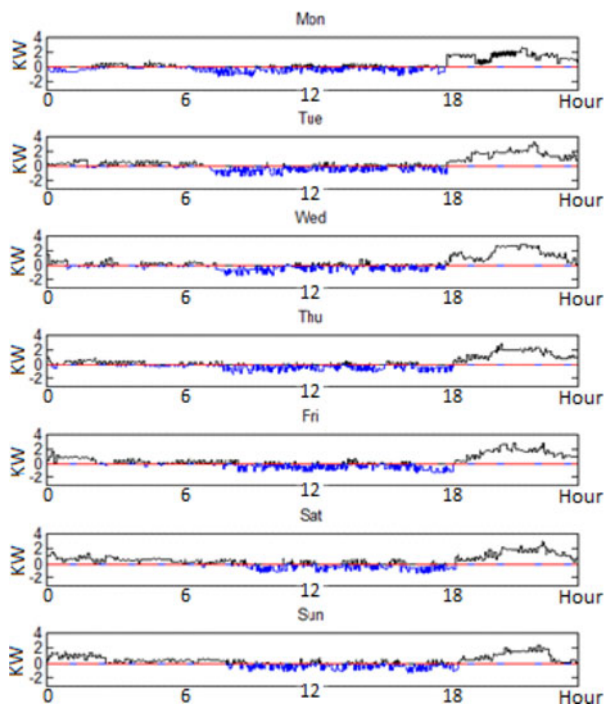

Figure 10. SEH output after applying EMS.

Table 3. Summarised information of the SEH before and after applying EMS.

Consumption (kW h)	No EMS	135.90
	EMS	117.02
Natural gas (therm)	No EMS	0
	EMS	10.701
Generation (kW h)	No EMS	0
	EMS	48.117
Cost ($) [output of the J function (Equation (3))]	No EMS	24.145
	EMS	16.237

where $|A(s)|$ represents the number of actions and ε is the parameter, $0 \le \varepsilon \le 1$ which is the probability of exploration.

4. Simulation

In this section, simulation results have been depicted and performance assessment of the proposed algorithm is presented. A case study has been implemented in 'MATLAB

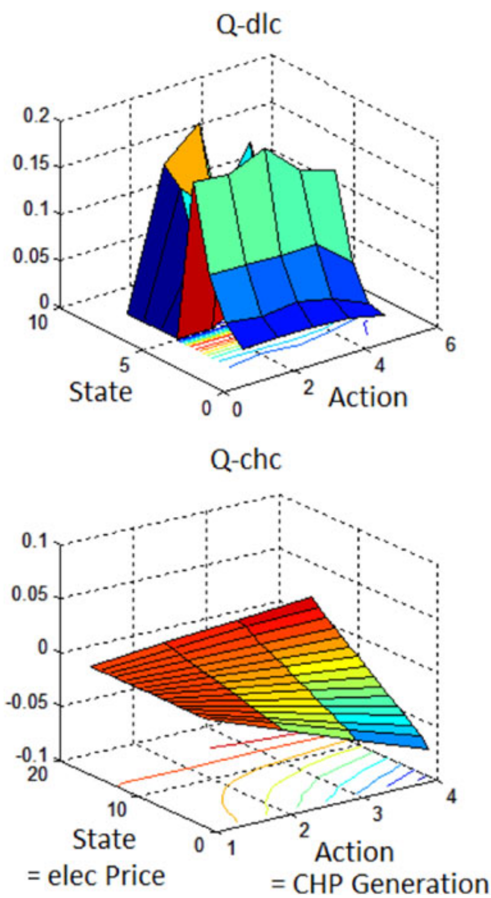

Figure 11. Q functions of four controllers.

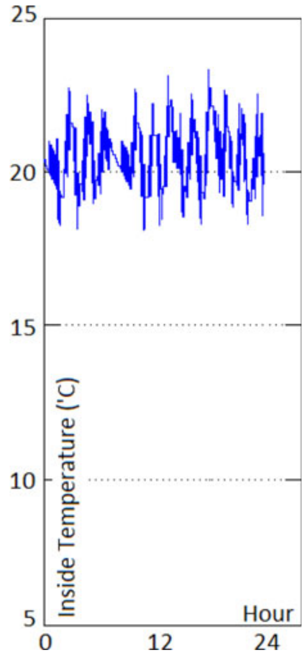

Figure 12. Inside temperature (°C).

Simulink' in real time and stochastic environment. It means the energy consumption of the SEH is not predetermined, and the price for every time slots is not calculated for the future times (Figure 7). The simulated SEH is shown in Figure 8.

We assume the average electricity price is 10.25 cents/kW h which is shown in Figure 7 and natural gas price is 112.02 cents/therm (1 therm = 2.74 m^3).[21] The first and second terms of dissatisfaction function prices P_d and P_c are determined by the consumers. In this article, we assume P_d is equal to 0.1 cent and P_c is equal to 0.5 cent.

The social cost of carbon dioxide (CO_2) is $21 per ton.[30] CO_2 emission from generation of electric power in the world in 2007 (α_{grid}) is 0.500 kg/kW h,[31] and CO_2 emission of CHP to generate heating and electric power (α_{CHP}) is 0.65714 kg/kW h.[32] The weather condition is derived from Tehran climate average of January presented in Table 1.[33]

PV, PHEV and CHP specifications are summarised in Table 2. We use an array of solar panel consisting of three solar modules 'BSP 80-12'.[34] The battery for PV is 'UB 1235-PS722'.[35] The 'lithium–manganese spinel/lithium–titanate' batteries for 20-mile PHEV [36] and a commercial micro CHP are used in this simulation.[37]

The simulation was run for seven consecutive weeks. SEH output before and after deploying EMS devices is shown in Figures 9 and 10, respectively.

After applying EMS in 7 weeks of natural gas and electricity consumption, electricity generation and bill cost of the SEH are presented and compared with the same system without EMS in Table 3.

Q functions of EMS learner in the final week are shown in Figure 11. In each step, Q functions determine what action should take place on a specific state (policy). The action that makes the Q function minimum is chosen in each step. For instance, the Q function of

CHP indicates that generation of CHP is high in prohibitive electric price condition and vice versa.

Inside temperature and PHEV's charging level are shown in Figures 12 and 13, respectively. For simplicity, we assume that PHEV is plugged in at 4 pm each day and is plugged out at 8 am.

The costs of electricity, natural gas and CO_2 in 7 weeks are demonstrated in Figure 14. In this regard, we can also see that natural gas consumption cost has been raised, total cost has plunged dramatically.

As shown here, with EMSs devices the peak load is reduced from 3.415 to 2.815 kW h (i.e. 17% less). The total energy cost reduces from \$24.145 to \$16.237 (i.e. 33% improvement) and CO_2 social cost has fallen to 0.889 (i.e. 53% less).

Note that a residence consumes less amount of electricity because it consumes natural gas energy, and CHP producing heating and does not need electricity for producing heat.

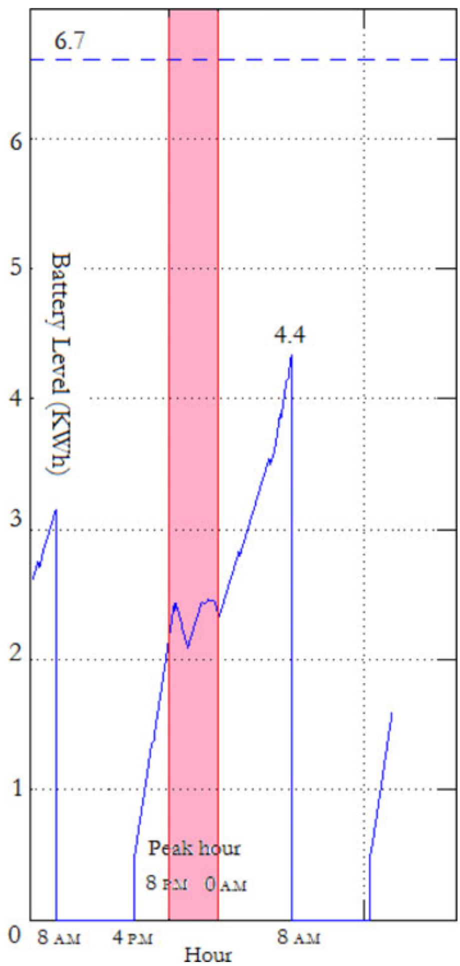

Figure 13. PHEV's charging level.

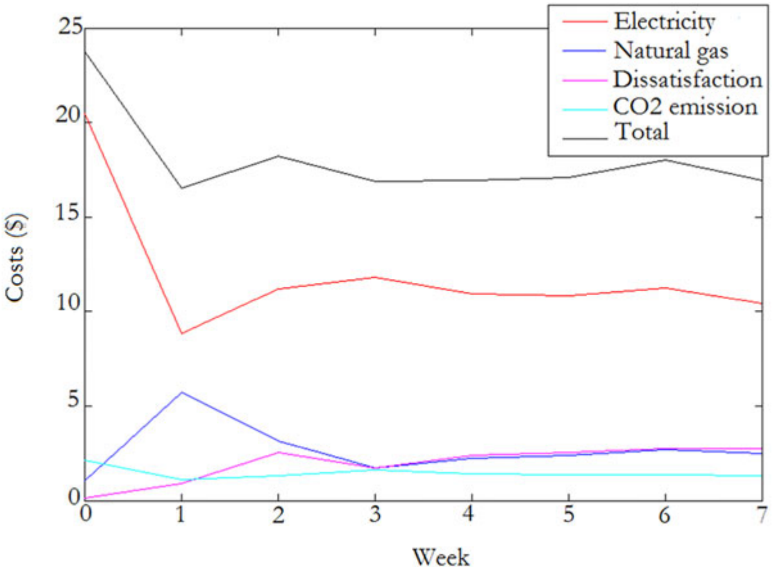

Figure 14. The costs in 7 consecutive weeks.

5. Conclusion

This paper introduces SEH as a new concept and proposes a practical method to find a near optimal solution in a very short time in order to be more efficient and simultaneously reduce electricity and natural gas bills. In our proposed method, we included dissatisfaction and CO_2 emission cost in total operational costs for more accuracy. To solve the model we applied RL algorithm. This algorithm does not require any details about dynamics of the system, hence it is more practicable. The algorithm converges to a near optimal solution swiftly and updates the solution during the time, hence we do not face any impediments by changing dynamics of a system. The simulation results confirm the efficiency of the model and algorithm and demonstrate up to 40% reduction in the electricity and gas cost. At the same time, peak load and CO_2 emission social cost are reduced by 17% and 50%, respectively.

References

[1] Harrison J. Micro Combined Heat & Power (CHP) for housing. 3rd International Conference on Sustainable Energy Technologies, Nottingham, UK; 2004. p. 1–6.
[2] Geidl M, Koeppel G, Favre-Perrod P, Klöckl B, Andersson G, Fröhlich K. Energy hubs for the futures. IEEE Power Energy Mag. 2007;5(1):24–30.
[3] Sheikhi A, Ranjbar AM, Oraee H. Financial analysis and optimal size and operation for a multicarrier energy system. Energy Buildings. 2012;48:71–78.
[4] Sheikhi A, Mozafari B, Ranjbar AM. CHP optimized selection methodology for a multi-carrier energy system. IEEE PowerTech, Trondheim, Norway; 2011. p. 1–7.
[5] Sheikhi A, Ranjbar AM, Safe F, Mahmoodi M. CHP optimized selection methodology for an energy hub system. In 10th International Conference on Environment and Electrical Engineering (EEEIC), Rome, Italy; 2011. p. 1–5.
[6] Parisio A, Del Vecchio C, Velotto G. Robust optimization of operations in energy hub. In 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC), Orlando, USA; 2011. p. 4943–4948.
[7] Houwing M, Negenborn R, DeSchtter B. Demand response with Micro-CHP systems. Proc IEEE. 2011;99(1):200–213.

[8] National Institute of Standards and Technology. NIST framework and roadmap for smart grid interoperability standards, release 1.0 (January 2010). Available from: http://www.nist.gov/public affairs/releases/upload/smartgrid interoperability final.pdf
[9] Samanthula BK, Chun H, Jiang W, McMillin BM. Secure and threshold-based power usage control in smart grid environments. Int J Parallel Emerg Distrib Syst. 2014;29(3):264–289.
[10] Fang X, Misra S, Xue G, Yang D. Smart grid – the new and improved power grid: a survey. IEEE Commun Surv Tutorials. 2012;14(4):944–980.
[11] Sheikhi A, Rayati M, Bahrami S, Ranjbar AM, Sattari S. A cloud computing framework on demand side management game in smart energy hubs. Int J Elec Power. 2015;64:1007–1016.
[12] Mohsenian-Rad A, Wong VWS, Jatskevich J, Schober R, Leon-Garcia A. Autonomous demand-side management based on game theoretic energy consumption scheduling for the future smart grid. IEEE Trans Smart Grid. 2010;1(3):320–331.
[13] He Q, Blum RS. New hypothesis testing-based rapid change detection for power grid system monitoring. Int J Parallel Emerg Distrib Syst. 2014;29(3):239–263.
[14] Li D, Aung Z, Williams JR, Sanchez A. No peeking: privacy-preserving demand response system in smart grids. Int J Parallel Emerg Distrib Syst. 2014;29(3):290–315.
[15] Oviedo Raul J. Martinez, Fan Zhong, Gormus Sedat, Kulkarni Parag, Kaleshi Dritan. Residential energy demand management in smart grids. In Transmission and Distribution Conference and Exposition (T&D), Orlando, USA; 2012. p. 1–8.
[16] Erol-Kantarci Melike, Sarker Jahangir H, Mouftah Hussein T. Communication-based plug-in hybrid electrical vehicle load management in the smart grid. In Computers and Communications (ISCC), Kerkyra, Greece; 2011. p. 404–409.
[17] Mohsenian-Rad A-H, Leon-Garcia A. Optimal residential load control with price prediction in real-time electricity pricing environments. IEEE Trans Smart Grid. 2011;1(2):120–133.
[18] Ibars, Christian, Navarro Monica, Giupponi Lorenza. Distributed demand management in smart grid with a congestion game. In First IEEE International Conference on Smart Grid Communications (SmartGridComm), Gaithersburg, USA; 2010. p. 495–500.
[19] O'Neill Daniel, Levorato Marco, Goldsmith Andrea, Mitra Urbashi. Residential demand response using reinforcement learning. In First IEEE International Conference on Smart Grid Communications (SmartGridComm), Gaithersburg, USA; 2010. p. 409–414.
[20] Chehreghani Bozchalui M, Hashmi A, Hassen H, Cañizares CA, Bhattacharya K. Optimal operation of residential energy hubs in smart grids. IEEE Trans Smart Grid. 2012;3(4):1755–1766.
[21] US Department of Energy. Buildings Energy Data Book, Annual Energy Outlook 2012 Early Release, January 2012. Available from: http://buildingsdatabook.eren.doe.gov/TableView.aspx?table=2.1.8
[22] Song G, Ding F, Zhang W, Song A. A wireless power outlet system for smart home. IEEE Trans Consum Electron. 2008;54(4):1688–1691.
[23] Rodríguez M, Gayoso JA. Degrees of freedom analysis for process control. 16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering, Madrid, Spain; 2006. p. 1489–1494.
[24] Han D-M, Lin J-H. Smart home energy management system using IEEE 802.15.4 and ZigBee. IEEE Trans Consum Electron. 2010;56(3):1403–1410.
[25] Matvoz D, Maksic M. Comparison of LED and CFL lamps and their impact on electric power network. In IEEE 15th International Conference on Harmonics and Quality of Power (ICHQP), Hong Kong; 2012. p. 320–327.
[26] Sutton Richard S. Introduction: The challenge of reinforcement learning. Reinforcement Learning. 1992;8(3):225–227.
[27] Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. J Artif Intell Res. 1996;4:237–285.
[28] Singh Satinder, Jaakkola Tommi, Littman Michael L, Szepesvári Csaba. Convergence results for single-step on-policy reinforcement-learning algorithms. Mach Learn, 2000;38(3):287–308.
[29] Sutton RS, Barto AG. Reinforcement learning: an introduction. London: MIT Press; 2005.
[30] Ackerman, Frank, Stanton Elizabeth A. The social cost of carbon. Economics for Equity and Environment (E3 Network). Available from: http://www.e3network.org/papers/SocialCostOfCarbon_SEI_20100401; 2010.

[31] European Environment Agency, CO_2 emission of each kW h. Available from: http://www.eea.europa.eu/dataandmaps/figures/co2-emissions-per-kwh-of ; 2007.

[32] Hedman Bruce A, Hampson Anne C. Fuel and CO_2 emissions savings calculation methodology for combined heat and power (CHP) systems. ASHRAE Trans. 2011;117(1):961–965.

[33] climatemps, Historical Weather For 2012 in Tehran. Available from: http://www.tehran.climatemps.com/; 2013.

[34] powerup co, Product information sheet solar module BSP80-12. Available from: http://www.powerupco.com; 2012.

[35] Battery stuff, Product information sheet AGM® battery UB12350-D5722. Available from: http://www.batterystuff.com/batteries/upc-telecom/UB12350-D5722.html; 2012.

[36] Nelson Paul, Amine Khalil, Rousseau Aymeric. Advanced Lithium-Ion batteries for plug-in hybrid-electric vehicles. Lithium Batteries: Research, Technology and Applications; 2008. p. 203–210.

[37] Viessmann co, Micro CHP boiler Vitotwin 300-W: the boiler that generates power. Available from: http://www.viessmann.de; 2012.