Multiagent-based simulation of simultaneous electricity market auctions in restructured environment

Mohammad FARSHAD*
Department of Electrical Engineering, Minoodasht Branch, Islamic Azad University, Minoodasht, Iran

Received: 17.05.2013 • Accepted/Published Online: 11.09.2013 • Printed: 31.12.2015

Abstract: In the restructured environment of the power industry, various commodities such as energy and operating reserves may be provided through simultaneous auctions. Prediction of market players' behavior in the auctions and simulation of the markets' environment can assist market decision-makers in evaluating specific policies before enforcing them in the real environment. Considering effects of the energy and varieties of reserve markets and also their interactions in the simulations is of high importance, which leads to more realistic simulation results. In this paper, an approach based on a multiagent system is proposed for simulating the simultaneous energy, spinning reserve, and replacement reserve market auctions, in which bidding of each agent is carried out based on a reinforcement learning algorithm. The proposed method is applied on a sample system, through which the impacts of considering lost opportunity cost in the payment model are examined.

Key words: Energy market, spinning reserve market, replacement reserve market, multiagent system, reinforcement learning

1. Introduction

Power industry restructuring and importance of electricity financial markets have motivated many researchers to conduct widespread investigations for accessing simulation tools for electricity markets. Useful information can be provided for decision-makers of electricity markets and also for players of those markets through a reality-based simulation [1]. Consequently, it will lead to development and improvement of the whole structure of electricity markets. Pivotal issues regarding power industry restructuring have been raised such as design and execution of auctions, payment calculation, financial terms in market environment, and strategic behavior of participants. One of the important duties of market decision-makers is to analyze various electricity markets in possible models and designs and also to select the best of them with respect to determined general goals. Prediction of market players' behavior in the auctions and simulation of markets' environment can help market decision-makers in selecting appropriate market structure design and rules. Each of the players must be appropriately modeled for simulating electricity markets. A common method for modeling each of the market players is to use intelligent agents. These intelligent agents will adopt appropriate strategies or make suitable decisions with respect to learning through gaining experience in the environment. In fact, an agent can control its decisions and actions according to the environment and can intelligently interact with the environment.

Various methods have been proposed for simulating electricity markets based on multiagent systems, some of which can be found in [2–18]. In the majority of the efforts made, only the energy market was taken

*Correspondence: mohammad.farshad@alumni.um.ac.ir
into account in the simulations and studies. It is important to consider the fact that there are other markets such as various ancillary service markets as well as the energy market, whose design is crucial in addition to designing the energy market. Operating reserves as the most important ancillary services are procured to provide appropriate reliability of the system. Electric generators are devices that have the capability of providing energy, reserve capacity, and other kinds of products at the same time. Hence, offers to the operating reserve market auctions may come from the generators, which may offer their products in the energy market, too.

This dependence on the supply side of the energy and reserve markets may result in interactions between these markets. In this respect, design of the energy and reserve markets should be carried out in a coordinated way [19,20]. Therefore, considering interactions of the energy market and varieties of reserve markets in simulations is of high importance, which leads to more realistic simulation results. In some cases, such as [2,14,16–18], effects of ancillary services were included in multiagent system-based studies and simulations; however, there were not flexible and detailed simulation methods to do studies and analyses of different possible designs in the reserve markets in association with the energy market.

Besides all existing market simulation methods, the overall aim of this paper is to propose a flexible and detailed multiagent-based simulation approach that can be used to do analyses of different possible designs and models in the simultaneous auctions of energy and reserve markets. In the multiagent system, selection of action for each agent can be done with different methods and algorithms. The simulation approach proposed in this paper is established based on utilizing a reinforcement learning-based multiagent system [2–4,7–10,13–15], in which each agent simulates the behavior of one of the electricity market participants. In the proposed multiagent system, each agent performs bidding in the simultaneous auctions of energy and various reserve markets based on Q-learning, which is a kind of reinforcement learning algorithm.

In this paper, for clear discussion and for better investigation of the proposed method, a main framework is considered. It should be noted that considering this framework will not violate the generality of the proposed simulation method. In fact, the proposed approach can be adaptable for a broad range of studies, taking into account some related special considerations. In this effort, two operating reserve types are considered, including the spinning reserve (SR) and the replacement reserve (RR). The SR requirements, which should be available within 10 min, can be supplied by only online generators. The RR requirements, which should be available within 60 min, can be provided by both online and offline generators. In this paper, the energy, SR, and RR requirements are procured by an independent system operator (ISO) through pool-based day-ahead markets. The pricing rule of these markets is considered to be uniform. Accordingly, the generators submit their bids for each hour of the operating day to the market auctions. Then the ISO minimizes the social cost of energy and reserve provision taking into account the submitted bids and the relevant constraints and requirements. Finally, the market participants will be informed of markets’ clearing prices and respective assigned capacities in each of the markets for each hour of the operating day.

In the following text, details regarding the proposed Q-learning-based multiagent system for simulating the simultaneous auctions of energy, SR, and RR markets are presented. Then the proposed approach is applied on a sample system with 17 generators, through which impact of considering the lost opportunity cost (LOC) [21] in the model of payments to auctions’ participants is examined. It is worth noting that in the sample system under study, the demand substitution possibility [22] is also considered.
2. Multiagent-based simulation

In the multiagent system considered for simulating the energy, SR, and RR markets, each of the agents, which are indicative of the market participants, submit their bids based on a reinforcement learning algorithm to the auctions. Figure 1 shows the general schematic view of the multiagent system structure considered for simulating the energy and reserve markets. In the following, first a brief description of the reinforcement learning algorithm in use is presented. Then the modeling procedures of the two basic elements, environment and agent, are presented, which are essential in implementing the reinforcement learning algorithm.

2.1. Reinforcement learning

In reinforcement learning, an agent learns which action suits which state. In fact, reinforcement learning is a mapping of state space to action space in such a way that the reward is maximized in the long term. In this algorithm, the agent is not told what action is appropriate to take (unsupervised learning) and it should find an action that results in a higher reward through interaction with the environment. Q-learning, as one of the reinforcement learning algorithms, is very common in multiagent-based simulations because of its simplicity and high efficiency. This algorithm, which is categorized as a temporal difference learning method [23], is applicable for solving Markovian decision problems. The Q-learning algorithm is appropriate for decision-making problems in repetitive games in which other players’ identities are unknown, as well as strategic bidding in electricity market auctions [3,24]. Figure 2 shows a schematic view of an agent’s interaction with its environment in the Q-learning algorithm.

---

**Figure 1.** Schematic view of the multiagent system structure.

**Figure 2.** Schematic view of an agent’s interaction with its environment.

For presentation of mathematical formulation of the Q-learning algorithm, imagine that an agent is interacting with its environment in discrete time stages. The finite series of states and actions for this agent are $S=\{s_1, s_2, \ldots, s_g\}$ and $A=\{a_1, a_2, \ldots, a_w\}$, respectively. At each time stage $t$, the agent senses its environment’s current state, $s_t \in S$, and selects the action $a_t \in A$ based on that input. Due to the selection of this action, the agent receives the reward $r_{t+1}$, and the state of the environment changes to a new one.
The agent’s target is to find the optimal policy $\pi^*(s) \in A$ for each state $s$, in a way that the overall gained reward is maximized in the long term [3,23,25]. The Q-learning algorithm finds the optimal policy via estimating optimal amounts of $Q$ value for each state–action pair, $Q^*(s,a)$. In fact, the optimal strategy can be defined as [23]:

$$
\pi^*(s) = \arg \max_a (Q^*(s,a)).
$$

(1)

The Q-learning algorithm finds values of $Q^*(s,a)$ through a recursive and repetitive procedure and based on available information $s_{t+1}$, $a_t$, $s_t$, $r_{t+1}$. The updating process in the algorithm is defined as [23]:

$$
Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha (r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)),
$$

(2)

where $\alpha (0 < \alpha \leq 1)$ is the learning rate and $\gamma (0 \leq \gamma < 1)$ is the discount factor. The learning rate $\alpha$ reflects the learning behavior of agent. If the value of $\alpha$ for an agent is considered to be greater, then the effect of the newly acquired information including the destination state and the gained reward will be more and the effect of the old $Q$ value will be less in the $Q$ value updating stage. The discount factor $\gamma$ reflects the farsighted behavior of the agent. Considering greater values of $\gamma$ will direct the agent to strive for a long-term high reward. On the contrary, smaller values of $\gamma$ will intensify the opportunistic behavior of the agent.

2.2. Environment simulation

The environment for each agent (supplier) includes conditions and rules of the markets and also the behavior of other suppliers in the markets. As shown in Figure 2, each agent receives two input signals including the state signal and the reward signal from the environment. Since the market pricing is done in a uniform manner, the environment’s state at each hour is determined according to the energy, SR, and RR market clearing prices. For avoiding the curse of dimensionality, a limited number of states between the minimum and maximum possible prices in the energy, SR, and RR markets are considered. Figure 3 demonstrates a schematic view of the state space at each hour.

![Figure 3. Schematic view of the state space at each hour.](image)

The reward of each agent at each hour is calculated based on its profit at that hour. The reward of each agent at hour $h$ for action $a$ in state $s$ can be calculated as [3]:

$$
r_h(s,a) = r(h) \times \left(\frac{AUR(h)}{TU(h)}\right)^b,
$$

(3)

2243
where \( r(h) \) is the agent’s received payment at hour \( h \) minus the supply costs at that hour; \( TUR(h) \) and \( AUR(h) \) stand for the target and actual utilization rate of the agent’s capacity at hour \( h \), respectively; and \( b \) indicates the significance of the agent’s capacity utilization rate.

### 2.3. Agents’ bidding strategy

The agents’ bidding strategy is designed based on the Q-learning algorithm. Here it is assumed that market participants offer all their maximum generation capability and there is no concealment of generation capacity. Each agent makes a decision for its bidding in the energy, SR, and RR markets’ auctions for each hour. For avoiding the curse of dimensionality, the action space for each agent is divided into parts between the agent’s supply costs and the maximum possible prices in the energy, SR, and RR markets. It is worth noting that taking a specific action means selecting a random price in the related span in the action space. Figure 4 demonstrates a schematic view of the action space for each agent at each hour.

Based on the Q-learning algorithm [3,23], three stages for the initial learning and main bidding process of each agent are state determination, action selection, and updating state–action values. At the beginning of each trading day, each agent senses the state of that day based on the energy, SR, and RR markets’ clearing prices in the last 24 h. Next, with respect to the state–action value at each hour, \( Q(s, a) \), each agent selects an action for that hour based on the \( \varepsilon \)-greedy strategy adopted for the purpose of balancing between the exploration and the greedy selection. Then, at the end of each trading day, when the amount of allocated power to each agent and also the market clearing prices at each hour are determined, each agent updates state–action values, \( Q(s, a) \), using Eq. (2).

It is worth noting that in the \( \varepsilon \)-greedy selection strategy, the action related to the best value of \( Q \) may be selected with the probability of \( 1 - \varepsilon \), and random action selection may be carried out without considering the values of \( Q \) with the probability of \( \varepsilon \).

In the initial learning process, the above-mentioned three stages are carried out for several trading days to obtain appropriate initial state–action values, \( Q(s, a) \), for each of the agents. It is worth noting that during the initial learning process, the value of learning rate \( \alpha \) for each agent (used in Eq. (2)) at hour \( h \) and day \( d \) for action \( a \) in state \( s \) is calculated as [25]:

\[
\alpha_{h,d}(s, a) = \left[\theta_{h,d}(s, a)\right]^{-1}, \tag{4}
\]

where \( \theta_{h,d}(s, a) \) equals the number of times the state–action pair \( (s, a) \) at hour \( h \) is met until the trading day \( d \). In the main bidding process, the state–action values, \( Q(s, a) \), which are acquired in the initial learning
process, are used as initial start values. It is worth noting that after the initial learning and in the main bidding process, a predetermined fixed value of learning rate $\alpha$ is used for each agent.

3. Numerical studies

3.1. Sample system specifications

The specifications of generators in the study system, considering a variety of generator sizes and supply costs, are provided in Table 1. The seasonal load forecasted by ISO for 24 h a day is shown in Figure 5. The SR and RR requirements at each hour are considered to be 15% of the forecasted load at that hour.

<table>
<thead>
<tr>
<th>Gen. no.</th>
<th>$P_{\text{min}}$ (MW)</th>
<th>$P_{\text{max}}$ (MW)</th>
<th>$SR_{\text{max}}$ (MW)</th>
<th>$RR_{\text{max}}$ (MW)</th>
<th>Energy cost ($/\text{MWh}$)</th>
<th>SR cost ($/\text{MW}$)</th>
<th>RR cost ($/\text{MW}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150</td>
<td>440</td>
<td>100</td>
<td>300</td>
<td>20.0</td>
<td>7.5</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>440</td>
<td>100</td>
<td>300</td>
<td>22.0</td>
<td>4.5</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>300</td>
<td>100</td>
<td>300</td>
<td>20.0</td>
<td>5.0</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>250</td>
<td>100</td>
<td>250</td>
<td>24.0</td>
<td>4.0</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>80</td>
<td>250</td>
<td>140</td>
<td>250</td>
<td>28.0</td>
<td>7.5</td>
<td>2.5</td>
</tr>
<tr>
<td>6</td>
<td>80</td>
<td>200</td>
<td>100</td>
<td>200</td>
<td>22.0</td>
<td>5.0</td>
<td>3.0</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>200</td>
<td>100</td>
<td>200</td>
<td>29.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>8</td>
<td>65</td>
<td>200</td>
<td>135</td>
<td>200</td>
<td>26.0</td>
<td>5.5</td>
<td>4.5</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>150</td>
<td>100</td>
<td>150</td>
<td>31.0</td>
<td>3.5</td>
<td>3.0</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>150</td>
<td>100</td>
<td>150</td>
<td>33.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>150</td>
<td>100</td>
<td>150</td>
<td>29.0</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>12</td>
<td>35</td>
<td>100</td>
<td>65</td>
<td>100</td>
<td>35.0</td>
<td>2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>13</td>
<td>35</td>
<td>100</td>
<td>65</td>
<td>100</td>
<td>36.0</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>14</td>
<td>35</td>
<td>100</td>
<td>65</td>
<td>100</td>
<td>31.0</td>
<td>4.0</td>
<td>3.5</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>50</td>
<td>35</td>
<td>50</td>
<td>44.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>50</td>
<td>35</td>
<td>50</td>
<td>47.0</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>17</td>
<td>15</td>
<td>50</td>
<td>35</td>
<td>50</td>
<td>36.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Figure 5. Seasonal load forecasted by ISO for 24 h a day.

The minimum energy, SR, and RR prices for the system under study are considered to be zero. The maximum prices in the energy, SR, and RR markets are taken to be 150 $/\text{MWh}, 100 $/\text{MW}, and 100 $/\text{MW}, respectively. The average costs of energy, SR, and RR supply for each of the agents are available in Table 1. Based on what is mentioned in Sections 2.2 and 2.3 for determining the state and action spaces, the price spaces
in the energy, SR, and RR markets are divided into five spans between the specified minimum and maximum prices; furthermore, five spans between the related supply costs and the maximum prices are considered for agents' bidding in the energy, SR, and RR markets. In other words, the considered parameters in Figures 3 and 4 are taken to be as follows: \( n = m = p = 5 \) and \( i = j = k = 5 \).

Finally, based on the available knowledge about the suppliers, the behavioral characteristics of agents should be determined by the parameters introduced in Section 2 including \( \alpha, \gamma, b, TUR, \) and \( \varepsilon \). For the agents of the sample system, a variety of values for \( \alpha, \gamma, b, \) and \( TUR \) are considered. The exploration probability \( \varepsilon \) is considered to be the same for all agents. The parameter values for the participant agents are presented in Table 2. The values of parameters except for \( \alpha \) are used in both the initial learning process and the main bidding process. The values of \( \alpha \) presented in Table 2 are used only in the main bidding process of agents, while in the initial learning process, the calculations are done based on Eq. (4).

### Table 2. Parameter values for the participant agents in the system under study.

<table>
<thead>
<tr>
<th>Gen. no.</th>
<th>Learning rate ( \alpha )</th>
<th>Discount factor ( \gamma )</th>
<th>Significance of utilization rateb</th>
<th>Target utilization rate ( TUR )</th>
<th>Exploration probability ( \varepsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7</td>
<td>0.1</td>
<td>2.0</td>
<td>0.9</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>0.3</td>
<td>2.0</td>
<td>0.9</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
<td>2.0</td>
<td>0.8</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>0.7</td>
<td>0.1</td>
<td>2.0</td>
<td>0.8</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>0.4</td>
<td>0.3</td>
<td>2.0</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
<td>0.5</td>
<td>1.5</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>7</td>
<td>0.1</td>
<td>0.5</td>
<td>1.5</td>
<td>0.9</td>
<td>0.25</td>
</tr>
<tr>
<td>8</td>
<td>0.4</td>
<td>0.3</td>
<td>1.5</td>
<td>0.9</td>
<td>0.25</td>
</tr>
<tr>
<td>9</td>
<td>0.7</td>
<td>0.1</td>
<td>1.5</td>
<td>0.9</td>
<td>0.25</td>
</tr>
<tr>
<td>10</td>
<td>0.4</td>
<td>0.3</td>
<td>1.5</td>
<td>0.8</td>
<td>0.25</td>
</tr>
<tr>
<td>11</td>
<td>0.1</td>
<td>0.5</td>
<td>1.5</td>
<td>0.8</td>
<td>0.25</td>
</tr>
<tr>
<td>12</td>
<td>0.4</td>
<td>0.3</td>
<td>1.0</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>13</td>
<td>0.7</td>
<td>0.1</td>
<td>1.0</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>14</td>
<td>0.1</td>
<td>0.5</td>
<td>1.0</td>
<td>0.9</td>
<td>0.25</td>
</tr>
<tr>
<td>15</td>
<td>0.4</td>
<td>0.3</td>
<td>1.0</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>16</td>
<td>0.7</td>
<td>0.1</td>
<td>1.0</td>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>17</td>
<td>0.1</td>
<td>0.5</td>
<td>1.0</td>
<td>0.9</td>
<td>0.25</td>
</tr>
</tbody>
</table>

### 3.2. Models under analysis in the sample system

The proposed multiagent system is carried out on the sample system to analyze effects of considering the LOC in the payment model. In this regard, two payment models are compared under similar conditions; in the first one, the winners of reserve auctions receive only the availability payment (Model \( A \)) [21], and in the second one, the winners may receive both the availability and LOC payments (model \( A+L \)) [21]. The availability payment is paid proportional to the amount of reserve capacity assigned to the winner, and the LOC is the generator’s lost profit in the energy market due to reserve capacity assigned to it. It is worth noting that for both above-mentioned payment models, the demand substitution possibility [22] is considered in the simultaneous allocation of energy, SR, and RR. Quality of the reserve commodities is usually determined with respect to the responding time and reliability. In this manner, SR has higher quality than RR. Demand substitution possibility means that the amount of the system’s need for each kind of reserve can be substituted regarding the quality of reserve product. For instance, the system RR requirement can be reduced and the same reduced amount can be added to the system SR requirement.
The formulations of the allocation problem for the specific models under analysis in the sample system, Model (A) and Model (A+L), are required to be used in the simulations. Model (A) and Model (A+L) can be examined through considering the relevant distinct features in the allocation cooptimization problem formulation and also in the calculation of participants’ payments. As mentioned in the introduction, the ISO minimizes the social cost of energy and reserve provision considering the submitted bids and subject to the relevant constraints and requirements. It is worth noting that for more simplicity and faster processing, the time-dependent constraints of generators such as the minimum up/down time constraints are not activated here; thus, the allocation cooptimization problem can be solved disjointedly for each hour of the day. Under Model (A), the energy, SR, and RR allocation cooptimization problem for an hour of the operating day can be expressed as in Eqs. (5a)–(5h), in which Eq. (5a) presents the objective function and Eqs. (5b)–(5h) present the constraints:

\[
\min \sum_i (g_i P_i \cdot p_{ei} + g_i S_{Ri} \cdot p_{sr_i} + RR_i \cdot p_{rr_i}),
\]

\[
g_i P_i + g_i S_{Ri} + RR_i \leq P_i^{\text{max}},
\]

\[
P_i^{\text{min}} \leq P_i,
\]

\[
0 \leq S_{Ri} \leq S_{Ri}^{\text{max}},
\]

\[
0 \leq RR_i \leq RR_i^{\text{max}},
\]

\[
\sum_i g_i P_i = P^\text{Load},
\]

\[
\sum_i g_i S_{Ri} \geq S_R^{\text{req}},
\]

\[
\sum_i RR_i + \sum_i g_i S_{Ri} = RR^{\text{req}} + S_{R}^{\text{req}},
\]

where \( P_i, S_{Ri}, \) and \( RR_i \) stand for the output power, SR capacity, and RR capacity assigned to generator \( i \), respectively; \( p_{ei}, p_{sr_i}, \) and \( p_{rr_i} \) are the offered price of generator \( i \) for energy, SR, and RR, respectively; \( P_i^{\text{max}} \) and \( P_i^{\text{min}} \) indicate the maximum and minimum generation capacity of generator \( i \), respectively; \( S_{Ri}^{\text{max}} \) and \( RR_i^{\text{max}} \) show the maximum SR and RR capacity of generator \( i \), respectively; \( P^\text{Load} \) indicates the load demand; \( S_R^{\text{req}} \) and \( RR^{\text{req}} \) are the system SR and RR requirement, respectively; and \( g_i \) is a binary integer to specify whether generator \( i \) is committed in the energy market or not (1 = on, 0 = off).

Under Model (A+L), the energy, SR, and RR allocation cooptimization problem’s constraints are similar to those of Eqs. (5b)–(5h). Under this condition, the objective function of this cooptimization problem for an hour of the operating day can be expressed as:

\[
\min \sum_i (g_i P_i \cdot p_{ei} + g_i S_{Ri} \cdot p_{sr_i} + RR_i \cdot p_{rr_i} + LOC_i).
\]

The lost opportunity price (LOP) and LOC of generator \( i \) (\( LOP_i \) and \( LOC_i \)) in a market with the uniform clearing rule are determined as follows [21]:

\[
LOP_i = \begin{cases} 
\lambda - p_{ei}, & \lambda > p_{ei} \\
0, & \lambda \leq p_{ei}
\end{cases},
\]

2247
\[ LOC_i = \max \left\{ 0, LOP_i (\hat{P}_i - P_i) \right\} \]  

where \( \lambda \) stands for the energy market clearing price and \( \hat{P}_i \) is the output power of generator \( i \) obtained from solving the problem of Eq. (5) without considering the reserve requirements, i.e. considering \( SR^{Req} = RR^{Req} = 0 \).

It can be comprehended that there is a self-referential difficulty in the LOC calculation under Model \((A+L)\), since the price of \( \lambda \) is unknown until the optimization problem is solved. In this paper, the optimization problems are solved using a genetic algorithm (GA), which does not face self-referential trouble.

### 3.3. Analysis of payment models

Model \((A)\) and Model \((A+L)\) are compared in the sample system under study. In the initial learning process and to gain access appropriate initial state–action values, \( Q(s, a) \), agents’ bidding algorithms are carried out for 5000 electricity trading days under Model \((A)\) and Model \((A+L)\) separately.

In the main bidding process, the initial state–action values, \( Q(s, a) \), which are acquired in the initial learning process, are used as the start values. In this stage, agents’ bidding algorithms are conducted for 1000 electricity trading days under Model \((A)\) and Model \((A+L)\) separately. After execution of simulations for 1000 electricity trading days, the average clearing prices for the energy, SR, and RR markets are acquired for the two payment models. Figures 6–8 show the average clearing prices under Model \((A)\) and Model \((A+L)\) for the energy, SR, and RR markets, respectively.

As can be observed from Figure 6, the clearing prices of the energy market under Model \((A+L)\) are lower than those obtained under Model \((A)\). On the contrary, as can be seen in Figures 7 and 8, the clearing prices of SR and RR markets under Model \((A+L)\) are higher than those obtained under Model \((A)\). In the system under study, the market pricing rule is uniform and the payments to winners of market auctions are calculated based on the highest accepted bid. On the other hand, the demand in the energy market is usually higher than the demands in the reserve markets. In a market with a uniform pricing rule, the lowness of the energy market clearing price due to the highness of the demand in this market is preferred, since it may lead to a lower total
payment. Therefore, under Model (A+L), the price decline observed in the energy market promises a better condition despite the price increase observed in the SR and RR markets.

After execution of simulations for 1000 electricity trading days, the average payments to winners of markets are calculated for Model (A) and Model (A+L). Figures 9–12 show the average payments in the energy, SR, and RR markets and also the average LOC payments under the two payment models. The average total payments for Model (A) and Model (A+L), which are obtained from the sum of the LOC payments and the payments in the energy, SR, and RR markets, are presented in Figure 13, as well.

As can be observed from Figure 9, the payments in energy market under Model (A+L) are lower than when Model (A) is adopted. In contrast, considering Figures 10 and 11, it is obvious that the payments in the SR and RR markets under Model (A+L) are higher than when Model (A) is used. As is expected, the value of LOC for Model (A) equals zero, which is shown in Figure 12. Noticing Figure 13, it can be concluded that the
total payments under Model (A+L) are lower than when Model (A) is considered. Considering the forecasted load shown in Figure 5 and the results demonstrated in Figures 9–13, it can be concluded that the difference in the payments under Model (A) and Model (A+L) during the peak load hours, in which the value of LOC is higher, too, is considerably great.

![Average LOC payments under Model (A) and Model (A+L).](image1)

**Figure 12.** Average LOC payments under Model (A) and Model (A+L).

![Average total payments under Model (A) and Model (A+L).](image2)

**Figure 13.** Average total payments under Model (A) and Model (A+L).

### 3.4. Observations summary

Here, some key points that the decision-maker may deduce from the simulation results and comparison of Model (A) and Model (A+L) in the sample system are highlighted as follows:

I. The clearing prices of the energy market in a state in which the LOC is considered in the objective function of the allocation cooptimization problem and in the payment model (Model (A+L)) are lower than when the LOC is not considered (Model (A)).

II. The clearing prices of SR and RR markets under Model (A+L) are higher than those obtained under Model (A).

III. The payments in energy market under Model (A+L) are lower than those obtained under Model (A).

IV. The payments in SR and RR markets under Model (A+L) are higher than when Model (A) is considered.

V. The total payments under Model (A+L) are lower than when Model (A) is adopted. Moreover, the difference in the payments under the two payment models during the peak load hours is considerably great.

VI. It seems that in the sample system under study with the uniform settlement rule, where a lower energy market’s clearing price is preferred because of the higher demand, considering the LOC in the objective function of the allocation cooptimization problem and in the payment model has some merits.

The performed analysis, through which the effects of considering the LOC in the payment model are investigated, requires simultaneous consideration of the energy and operating reserve markets, which is provided.
by the proposed method. The proposed simulation approach is also suitable for a broad range of studies, such as comparing the different reserve market designs and pricing rules [21,22], investigating the possibility of price reversal [22], assessment of the effects of demand response [26], and interaction analysis between the markets [27].

4. Conclusion
Designing an appropriate structure and making effective rules for energy and operating reserve markets are of particular importance from technical and economic efficiency perspectives. In this paper, a flexible approach is proposed for simulation of the simultaneous auctions of energy and reserve markets based on a multiagent system and by implementing the Q-learning algorithm, through which varieties of possible market designs and rules can be analyzed. Significant results are obtained by implementing the proposed approach on a sample system and comparing the two payment models with and without considering the lost opportunity cost (LOC). In the system under study, considering the LOC in the payment model results in a modification of participants’ bidding behavior and also in a decrease in the total payment to them. Many market analyses, such as those done for investigating effects of considering the LOC in the payment model, require simulation methods in which the energy and operating reserve markets are considered concurrently. The proposed simulation approach facilitates performing such kinds of analyses through taking into account the interaction of the energy and a variety of reserve markets.

Acknowledgment
This article is the result of a research project supported by the Minoodasht Branch of Islamic Azad University. The author would like to acknowledge this support and contribution.

References


