

## تجزیه و تحلیل مدل پرداخت به برندگان مناقصات همزمان بازار انرژی و رزرو چرخشی بر پایه سیستم چند عاملی

محمد فرشاد      محمد باقر نقیبی سیستانی  
دانشجوی دکتری      استادیار  
گروه برق، دانشکده مهندسی، دانشگاه فردوسی مشهد  
مشهد، خراسان رضوی، ایران

واژه‌های کلیدی: بازار انرژی، بازار رزرو چرخشی، هزینه فرصت از دست رفته، یادگیری تقویتی، سیستم چند عاملی

### چکیده

در محیط تجدید ساختار یافته صنعت برق، انتخاب طرح‌ها و قوانین مناسب برای بازار انرژی و انواع سرویس‌های جانبی از دید تصمیم‌گیران بازار دارای اهمیت ویژه‌ای به لحاظ فنی و اقتصادی می‌باشد. انتخاب مدل پرداخت به برندگان مناقصه‌ها نیز یکی از موارد کلیدی در طراحی است که تاثیر مستقیم بر کارایی سیستم و راهبرد سیاست‌گذاری‌ها دارد. پیش‌بینی رفتار شرکت‌کنندگان در مناقصات بازار و شبیه‌سازی محیط بازار، در انتخاب مدل و قانون مناسب سودمند و کارا خواهد بود. در این مقاله، روشی مبتنی بر سیستم چند عاملی جهت شبیه‌سازی مناقصه همزمان بازار انرژی و رزرو چرخشی پیشنهاد شده است، که در آن قیمت‌دهی هر عامل بر اساس یک الگوریتم یادگیری تقویتی صورت می‌پذیرد. روش مذکور در خصوص یک سیستم نمونه با قانون تسویه یکنوخت اجرا شده و از این طریق تاثیر دو مدل پرداخت مختلف، با لحاظ هزینه فرصت از دست رفته و بدون آن، در رفتار قیمت‌دهی شرکت‌کنندگان بازارها و نتایج بازار مورد تجزیه و تحلیل قرار گرفته است.

### ۱- مقدمه:

یکی از وظایف مهم قانون‌گذاران و تصمیم‌گیران بازار برق، تجزیه و تحلیل بازار در مدل‌ها و طرح‌های ممکن و انتخاب بهترین آنها با توجه به سیاست‌گذاری‌ها و اهداف تعیین شده می‌باشد. شناخت کامل محیط بازار و شناخت ویژگی‌های رفتاری شرکت‌کنندگان بازار و پیش‌بینی رفتار قیمت‌دهی از ملزومات این امر محسوب می‌شود.

در محیط صنعت برق علاوه بر بازار انرژی، معمولاً بازارهای دیگری همچون بازار انواع سرویس‌های جانبی نیز وجود دارند که طراحی آنها به همراه طراحی بازار انرژی دارای اهمیت است. ظرفیت رزرو، به عنوان مهمترین عنصر جهت تامین قابلیت اطمینان مطلوب سیستم، یکی از انواع سرویس‌های جانبی است که در اکثر سیستم‌ها بصورت مشترک وجود دارد.

در اکثر سیستم‌های تجدید ساختار یافته برق مانند کالیفرنیا، نیویورک، PJM و New England، تامین سرویس‌های جانبی بر عهده بهره‌بردار مستقل سیستم (ISO) است [۱]. برای تامین

## بیست و پنجمین کنفرانس بین‌المللی برق

نظر قرار دارد و میزان مورد نیاز انواع کالاهای برق بصورت همزمان به فروشندگان تخصیص می‌یابد.

ژنراتورها تجهیزاتی هستند که توانایی عرضه انرژی، ظرفیت رزرو و دیگر انواع کالا را در یک زمان دارا هستند. از این رو ممکن است عرضه در بازار رزرو از طرف ژنراتورهایی باشد که در بازار انرژی نیز کالای خود را عرضه کرده‌اند. این وابستگی در سمت عرضه بازارهای انرژی و رزرو ممکن است باعث تاثیر متقابل این بازارها بر یکدیگر شود. اگر چه ممکن است قیمت رزرو با قیمت انرژی متفاوت باشد، ولی بازار رزرو و انرژی باید هماهنگ با یکدیگر طراحی گردند. [۶,۷]

علاوه بر دلایل فوق، یک ارتباط داخلی صریح بین بازار انرژی و بازار رزرو چرخشی وجود دارد که در مراجع [۸] و [۹] نیز به این ارتباط پرداخته شده است. لذا در این مقاله تخصیص همزمان و بهینه‌سازی توام در بازار انرژی و رزرو چرخشی مد نظر قرار گرفته است.

به لحاظ زمانی، مناقصه در بازار رزرو می‌تواند بصورت Forward، Day-Ahead، Real-Time بلند مدت و یا ترکیبی از آنها (مانند بازارهای New England و کالیفرنیا [۱۰,۱۱]) اجرا شود. در این مقاله بازه زمانی اجرای مناقصات انرژی و رزرو، Day-Ahead بوده و شرکت کنندگان در روز قبل برای هر ساعت روز بهره‌برداری، قیمت و مشخصات فنی مربوط به تامین انرژی و رزرو چرخشی را پیشنهاد می‌دهند.

برای انتخاب پیشنهادهای برنده در بازار می‌توان دو نوع تابع هدف در نظر گرفت: [۱]

الف) هزینه اجتماعی<sup>۲</sup>

ب) هزینه تامین<sup>۳</sup>

در تابع هدف (الف)، مینیمم سازی بر اساس پیشنهادات قیمت ارائه شده توسط شرکت کنندگان و ظرفیت پذیرفته شده آنها صورت می‌پذیرد. در تابع هدف (ب)، مینیمم سازی بر اساس قیمت پرداختی به برندگان و ظرفیت پذیرفته شده آنها صورت می‌گیرد. در بازاری با قانون تسویه تمایزی<sup>۴</sup>، نتایج

رزرو در محیط رقابتی صنعت برق روش‌ها و مدل‌های مختلفی وجود دارد، ISO می‌تواند برای تامین رزرو موردنیاز سیستم بازاری با مدل Pool راه‌اندازی کند (مانند کالیفرنیا، نیویورک، PJM و New England) و یا تحت قراردادهای دو طرفه با تولید کنندگان این میزان رزرو را خریداری نماید (مانند اسلوانی، استرالیا و کشورهای شمال اروپا) [۲,۳]. امروزه تامین رزرو و دیگر خدمات جانبی از طریق بازار رقابتی مورد توجه قرار گرفته است، زیرا رقابت می‌تواند به افزایش سودمندی و بهره‌وری کالا، شفافیت قیمت و رضایت تولیدکننده و مصرف کننده کمک نماید. بازار رقابتی رزرو این اجازه را به شرکت کنندگان می‌دهد تا بطور مناسب بین هزینه و سود حاصل از تامین رزرو تعادل ایجاد کرده و ارزش واقعی کالای رزرو را مشخص نمایند [۴]. در این مقاله تامین رزرو به مانند تامین انرژی موردنیاز مصرف کننده از طریق یک بازار رقابتی Pool که توسط ISO اجرا می‌شود، مد نظر قرار گرفته است.

در هر سیستم با توجه به مشخصات واحدهای تولیدی آن، انواع کالای رزرو قابل تامین که برای آن سیستم نیز مفید باشد، تعریف می‌گردند. ظرفیت رزرو چرخشی به عنوان یکی از مهمترین انواع رزرو در این مقاله مد نظر قرار گرفته است. رزرو چرخشی توسط واحدهای سنکرون با شبکه تامین می‌گردد. ظرفیت رزرو چرخشی تخصیص یافته به واحدها باید در مدت حداکثر ۱۰ دقیقه قابل دسترسی باشد [۵].

بطور کلی دو شکل تخصیص کالاهای برق، از جمله انرژی و انواع کالای رزرو، وجود دارد که عبارتند از تخصیص پشت سرهم و تخصیص همزمان. در تخصیص پشت سرهم، معمولاً توالی تخصیص انواع کالای برق بر اساس اهمیت و کیفیت این کالاها تعیین می‌شود. مثلاً ابتدا مسئله بهینه‌سازی تخصیص انرژی حل می‌گردد، در مرحله بعد تخصیص در بازار رزرو چرخشی انجام می‌شود. در هر مرحله از تخصیص پشت سرهم، ظرفیت‌های تخصیص یافته به کالاهای دیگر در مراحل قبلی، از ظرفیت آماده حذف می‌گردند. در تخصیص همزمان، بهینه‌سازی توام<sup>۱</sup> هزینه انواع کالای انرژی و رزرو مد

2. Social Cost  
3. Procurement Cost  
4. Pay-as-Bid

1. Co-optimization

## بیست و پنجمین کنفرانس بین‌المللی برق

تجزیه و تحلیل قرار گیرد. برای تحقق این امر، استفاده از یک سیستم چند عاملی<sup>۲</sup> پیشنهاد شده است که در آن هر یک از عامل‌ها بر اساس یک الگوریتم یادگیری تقویتی<sup>۳</sup> در بازار انرژی و رزرو چرخشی قیمت‌دهی می‌کنند. برای بررسی‌های بیشتر، یک سیستم نمونه معرفی شده و روش پیشنهادی بر روی آن اجرا شده و در نهایت نتایج تجزیه و تحلیل رفتار شرکت‌کنندگان تحت دو مدل پرداخت مذکور آورده شده است. لازم به ذکر است، الگوریتم یادگیری تقویتی مورد استفاده، الگوریتم Q-Learning می‌باشد.

### ۲- الگوریتم Q-Learning:

در یادگیری تقویتی، عامل فرا می‌گیرد که در هر موقعیت و حالت چه عملی بهتر است. در واقع یادگیری تقویتی نگاهی از فضای موقعیت‌ها به فضای عمل‌ها می‌باشد، به نحوی که سیگنال پاداش ماکزیمم گردد. در یادگیری تقویتی مانند بسیاری از الگوریتم‌های یادگیری ماشین، به عامل گفته نمی‌شود که چه عملی را انتخاب کند (یادگیری بدون ناظر<sup>۴</sup>) و عامل باید در تقابل با محیط و با آزمایش عمل‌ها، عملی را که پاداش بیشتری در پی دارد، پیدا کند. [۱۵]

الگوریتم Q-Learning یک الگوریتم یادگیری تقویتی است که برای مسائل تصمیم‌گیری مارکوفی (MDPs)<sup>۵</sup> که اطلاعات کامل و روشنی از مدل آنها وجود ندارد، قابل استفاده می‌باشد. بر اساس این الگوریتم می‌توان استراتژی بهینه را در تعامل مستقیم با محیط و به صورت on-line بدست آورد. الگوریتم Q-Learning برای مسائل تصمیم‌گیری در بازی‌های تکراری که بقیه بازیگران ناشناخته هستند، مانند مسئله قیمت‌دهی استراتژیک در مناقصات بازار برق، مناسب خواهد بود. [۱۶، ۱۷]

نمونه‌هایی از به‌کارگیری الگوریتم Q-Learning در مسئله قیمت‌دهی استراتژیک در مناقصات بازار برق و شبیه‌سازی رفتار شرکت‌کنندگان بازار برق را می‌توان در مراجع [۲۱-۱۶] مشاهده نمود.

حاصل از این دو تابع هدف بر یکدیگر منطبق می‌باشند. روش مینیمم سازی هزینه تامین، کاملاً به قانون تسویه اعمالی در سیستم وابسته است. در بازاری با قانون پرداخت یکنواخت<sup>۱</sup>، نتایج مینیمم سازی هزینه اجتماعی لزوماً بر نتایج مینیمم سازی هزینه تامین منطبق نخواهد بود. [۱۲].

در این مقاله، قانون تسویه بازار به صورت یکنواخت در نظر گرفته شده است. تابع هدف مسئله بهینه سازی نیز، هزینه اجتماعی می‌باشد و مینیمم سازی بر اساس پیشنهادات قیمت ارائه شده توسط شرکت‌کنندگان و ظرفیت پذیرفته شده آنها صورت می‌پذیرد. لازم به ذکر است، مسئله بهینه سازی تخصیص، با ساده‌سازی و صرفنظر از محدودیت‌های وابسته به زمان، برای یک ساعت روز بهره‌برداری به صورت مجزا قابل حل خواهد بود.

بر اساس مرجع [۱۳]، در طراحی بازارهای رزرو ممکن است یکی از مدل‌های رایج پرداخت ذیل مورد استفاده قرار گیرد:

الف) ژنراتورها تنها هزینه آمادگی رزرو را دریافت کنند. (مدل (A))

ب) ژنراتورها تنها هزینه فرصت ازدست رفته را دریافت کنند. (مدل (L))

ج) ژنراتورها هزینه آمادگی رزرو و هزینه فرصت ازدست رفته را دریافت کنند. (مدل (A+L))

د) ژنراتورها هزینه آمادگی رزرو یا هزینه فرصت ازدست رفته را دریافت کنند. (مدل (A/L))

هزینه آمادگی رزرو مبلغی است که واحد به ازای مقدار ظرفیت رزرو تخصیص یافته به آن دریافت می‌کند. در مدل‌هایی که در آنها پرداخت هزینه آمادگی رزرو مدنظر قرار گرفته است، ژنراتورها باید قیمت آمادگی پیشنهادی خود را ارائه دهند. هزینه فرصت از دست رفته نیز در واقع برای جبران کاهش سود واحد در بازار انرژی بدلیل مطرح شدن قید تامین رزرو، پرداخت می‌شود [۱۴].

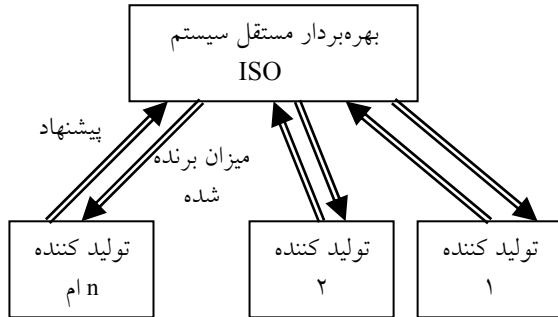
در این مقاله قصد بر آن است تا تاثیر دو مدل پرداخت (A) و (A+L) بر نتایج بازار و رفتار شرکت‌کنندگان بازار مورد

2. Multi-agent  
3. Reinforcement Learning  
4. Unsupervised Learning  
5. Markovian Decision Problems

1. Uniform

### بیست و پنجمین کنفرانس بین‌المللی برق

در رابطه فوق، پارامتر  $\alpha (0 < \alpha \leq 1)$  ضریب یادگیری و پارامتر  $\gamma (0 \leq \gamma \leq 1)$  ضریب مقیاس است.

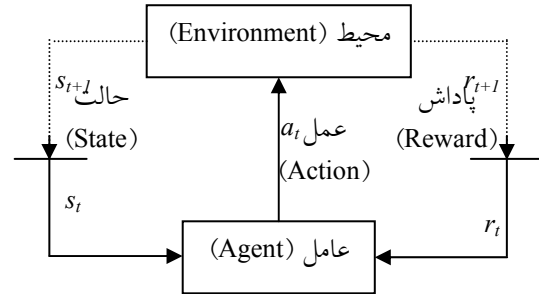


شکل ۲: شمایی کلی از ساختار سیستم چند عاملی

### ۳- سیستم چند عاملی:

همانطور که قبلاً نیز ذکر شد، در این مقاله قصد بر آن داریم تا رفتار بازیگران در مناقصه بازار انرژی و رزرو چرخشی را بر پایه سیستم چند عاملی مدل‌سازی کنیم. در این سیستم هر یک از عامل‌ها که مبین شرکت‌کنندگان بازار هستند، بر اساس الگوریتم یادگیری تقویتی در مناقصات قیمت‌دهی می‌کنند. همانطور که در شکل (۱) نیز مشخص است، برای شبیه‌سازی و اجرای الگوریتم یادگیری تقویتی، نیازمند آماده‌سازی دو رکن اساسی عامل و محیط خواهیم بود. شکل (۲)، شمایی کلی از ساختار سیستم چند عاملی موردنظر جهت شبیه‌سازی بازار انرژی و رزرو چرخشی را نمایش می‌دهد. در این سیستم، هر یک از شرکت‌کنندگان قیمت پیشنهادی و مشخصات فنی خود را برای هر ساعت روز بهره‌برداری برای تامین انرژی و رزرو چرخشی به ISO ارائه می‌دهند. ISO نیز پس از دریافت پیشنهادات شرکت‌کنندگان، مسئله بهینه‌سازی توام تخصیص انرژی و رزرو چرخشی را حل می‌کند و میزان ظرفیت برنده شده هر کدام را در بازار انرژی و رزرو چرخشی برای هر ساعت روز بهره‌برداری به آنها اعلام می‌کند. در سیستم مزبور، محیط برای هر عامل (تولیدکننده) شامل شرایط و قوانین بازار و همچنین رفتار سایر تولیدکنندگان در بازار می‌باشد. از آنجا که قانون تسویه بازار به صورت یکنواخت در نظر گرفته شده است، حالت در هر ساعت نیز قیمت بسته شدن بازار در آن ساعت

شکل (۱)، شمایی کلی از عملکرد عامل در تعامل با محیط در الگوریتم Q-Learning را نمایش می‌دهد.



شکل ۱: شمایی کلی از تعامل عامل با محیط

برای بیان فرمول‌بندی ریاضی الگوریتم Q-Learning، فرض کنید یک عامل در مراحل زمانی گسسته ( $t=1, 2, 3, \dots$ ) با محیط خود به صورت شکل (۱) در تعامل است. همچنین مجموعه متناهی حالت‌ها و عمل‌ها برای این عامل  $S = \{s_1, s_2, \dots, s_n\}$  و  $A = \{a_1, a_2, \dots, a_n\}$  می‌باشند. در هر مرحله زمانی  $t$ ، عامل حالت کنونی محیط خود را که  $s_t = s \in S$  می‌باشد، دریافت می‌کند و بر اساس آن یک عمل  $a_t = a \in A$  را انتخاب می‌نماید. در نتیجه انتخاب این عمل، عامل یک پاداش لحظه‌ای  $r_{t+1}$  را دریافت می‌کند و حالت محیط به حالت جدید  $s_{t+1} = s' \in S$  تغییر می‌یابد. هدف عامل یافتن استراتژی بهینه  $\pi^*(s) \in A$  برای هر حالت  $s$  می‌باشد، به گونه‌ای که کل پاداش دریافتی در بلندمدت ماکزیمم گردد. الگوریتم Q-Learning با تخمین مقادیر بهینه ارزش Q برای جفت حالت-عمل  $Q^*(s, a)$ ، استراتژی بهینه را پیدا می‌کند. در واقع استراتژی بهینه به صورت زیر بدست خواهد آمد: [۱۵، ۱۶]

$$\pi^*(s) = \arg \max_a (Q^*(s, a)) \quad (۱)$$

الگوریتم Q-Learning مقادیر  $Q^*(s, a)$  را از طریق یک روش بازگشتی و تکراری و بر اساس اطلاعات موجود  $s_{t+1}$ ،  $a_t$ ،  $s_t$  و  $r_{t+1}$  پیدا خواهد کرد. قانون به‌روز رسانی در این الگوریتم به صورت زیر می‌باشد: [۱۵، ۱۶]

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (۲)$$

### بیست و پنجمین کنفرانس بین المللی برق

$i$  ام در بازار انرژی (۰ یا ۱) خواهد بود.

قیود مسئله بهینه‌سازی توام تخصیص انرژی و رزرو چرخشی برای حالتی که ژنراتورها علاوه بر هزینه آمادگی، هزینه فرصت از دست رفته را نیز دریافت می‌کنند (مدل (A+L)) مشابه روابط (۲-۳) تا (۶-۳) خواهند بود. در این حالت تابع هدف این مسئله بهینه‌سازی برای هر ساعت روز بهره برداری به صورت رابطه (۴) قابل بیان است.

$$\text{Min} \sum_i (\alpha_i \cdot P_i \cdot \rho e_i + \alpha_i \cdot SR_i \cdot \rho sr_i + LOC_i) \quad (4)$$

قیمت فرصت از دست رفته (LOP) و هزینه فرصت از دست رفته (LOC) ژنراتور  $i$  ام، در بازاری با قانون تسویه یکنواخت، به ترتیب بر اساس روابط (۵) و (۶) بیان می‌شوند. [۱۳]

$$LOP_i = \begin{cases} \lambda - \rho e_i, & \lambda > \rho e_i \\ 0, & \lambda \leq \rho e_i \end{cases} \quad (5)$$

$$LOC_i = \max\{0, LOP_i(\hat{P}_i - P_i)\} \quad (6)$$

در روابط فوق،  $\lambda$  قیمت بسته شدن بازار انرژی و  $\hat{P}_i$  نتایج بدست آمده از حل مسئله بهینه‌سازی تخصیص تنهای انرژی و بدون لحاظ قیود تامین رزرو می‌باشد.

### ۲-۳- استراتژی قیمت‌دهی عامل‌ها:

استراتژی قیمت‌دهی عامل‌ها بر اساس الگوریتم Q-Learning پایه ریزی شده است. برای سادگی فرض می‌شود، شرکت‌کنندگان بازار تمامی ظرفیت ماکزیمم خود را به بازار پیشنهاد می‌دهند.

حالت‌های محیط در هر ساعت بر اساس قیمت بسته شدن بازار انرژی و رزرو چرخشی که به صورت عمومی در اختیار شرکت‌کنندگان قرار می‌گیرد، تعیین می‌شود. برای گریز از مشکل نفرین ابعاد، از گسسته‌سازی استفاده شده است، به این ترتیب که تعداد محدودی حالت بین قیمت کف و قیمت سقف بازار انرژی و رزرو چرخشی در نظر گرفته می‌شود.

هر عامل باید برای پیشنهاد قیمت خود در مناقصات بازار انرژی و رزرو چرخشی برای هر ساعت تصمیم‌گیری نماید. فضای عمل برای هر عامل به قسمت‌هایی بین هزینه تامین و

خواهد بود، که تمامی عامل‌ها از آن اطلاع پیدا خواهند کرد. برای هر عامل، پاداش لحظه‌ای در هر ساعت برابر میزان دریافتی وی خواهد بود.

### ۳-۱- مسئله بهینه‌سازی تخصیص:

در اینجا برای بیان مسئله تخصیص دو مدل پرداخت (A) و (A+L) را مد نظر قرار خواهیم داد.

در اینجا، ISO مسئله بهینه‌سازی توام تخصیص انرژی و رزرو چرخشی را بر اساس مینیمم سازی هزینه اجتماعی برای ۲۴ ساعت روز بهره‌برداری حل خواهد کرد. این مسئله برای حالتی که ژنراتورها تنها هزینه آمادگی رزرو را دریافت کنند (مدل (A)) برای هر ساعت روز بهره‌برداری در رابطه (۳) آمده است. رابطه (۱-۳) تابع هدف مسئله و روابط (۲-۳) تا (۶-۳) قیود مسئله بهینه‌سازی را نمایش می‌دهند. لازم به ذکر است، این مسئله با ساده‌سازی و صرفنظر از محدودیت‌هایی مانند حداقل زمان روشن و خاموش بودن واحدها و بطور کلی محدودیت‌های وابسته به زمان، برای هر ساعت روز بهره‌برداری به طور جداگانه قابل حل خواهد بود.

$$\text{Min} \sum_i (\alpha_i \cdot P_i \cdot \rho e_i + \alpha_i \cdot SR_i \cdot \rho sr_i) \quad (1-3)$$

$$P_i + SR_i \leq P_i^{\max} \quad (2-3)$$

$$P_i^{\min} \leq P_i \quad (3-3)$$

$$0 \leq SR_i \leq SR_i^{\max} \quad (4-3)$$

$$\sum_i \alpha_i P_i = P^{\text{Load}} \quad (5-3)$$

$$\sum_i \alpha_i SR_i = SR^{\text{Req}} \quad (6-3)$$

در رابطه (۳)،  $P_i$  و  $SR_i$  به ترتیب مقدار برنده شده واحد  $i$  ام در بازار انرژی و رزرو چرخشی و همچنین  $\rho e_i$  و  $\rho sr_i$  به ترتیب قیمت پیشنهادی واحد  $i$  ام در بازار انرژی و رزرو چرخشی می‌باشند.  $P_i^{\min}$  و  $P_i^{\max}$  نیز به ترتیب ماکزیمم و مینیمم توانایی تولید واحد  $i$  ام خواهد بود.  $SR_i^{\max}$  نیز بیانگر ماکزیمم توانایی تامین رزرو چرخشی واحد  $i$  ام است. در رابطه (۳)،  $P^{\text{Load}}$  و  $SR^{\text{Req}}$  به ترتیب میزان بار مصرفی سیستم و میزان ظرفیت رزرو چرخشی موردنیاز سیستم را مشخص می‌کنند. متغیر  $\alpha_i$  نیز نشان دهنده برنده شدن و یا نشدن واحد

### بیست و پنجمین کنفرانس بین المللی برق

چرخشی مربوط به ۲۴ ساعت گذشته، حالت هر ساعت آن روز مبادلات برق را دریافت می کند.

- **مرحله دوم (انتخاب عمل):** پس از تعیین حالت، هر عامل با توجه به جدول ارزش حالت-عمل  $(Q(s,a))$  در هر ساعت، عملی را برای هر ساعت بر اساس استراتژی  $\epsilon$ -greedy انتخاب می نماید. در استراتژی  $\epsilon$ -greedy، با احتمال  $\epsilon$  عمل مربوط به بهترین  $Q$  انتخاب می شود و با احتمال  $1-\epsilon$  بدون توجه به مقادیر  $Q$  انتخاب تصادفی عمل انجام خواهد شد. این امر به الگوریتم کمک می کند تا در فضای پاسخ بهتر کاوش کند و در واقع تعادلی بین کاوش و انتخاب حریصانه خواهد بود.

- **مرحله سوم (به روز رسانی جدول ارزش حالت-عمل):** در انتهای هر روز مبادلات برق، پس از آنکه میزان توان تخصیص یافته به هر عامل در بازار انرژی و رزرو و همچنین قیمت بسته شدن بازار در هر ساعت مشخص شد، هر عامل با توجه به پاداش لحظه ای بدست آمده از رابطه (۷)، مقادیر جدول ارزش حالت-عمل را از طریق رابطه (۲) به روز می نماید.

#### ۴- مطالعات عددی:

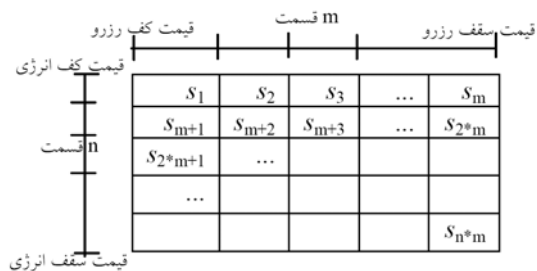
در اینجا، مشخصات یک سیستم نمونه بیان شده و سیستم چند عاملی مذکور در خصوص آن اجرا خواهد شد. در سناریوی اول از مدل پرداخت  $(A)$  و در سناریوی دوم از مدل پرداخت  $(A+L)$  استفاده خواهد شد و در نهایت، نتایج حاصل با یکدیگر مقایسه می شوند.

#### ۴-۱- معرفی سیستم نمونه:

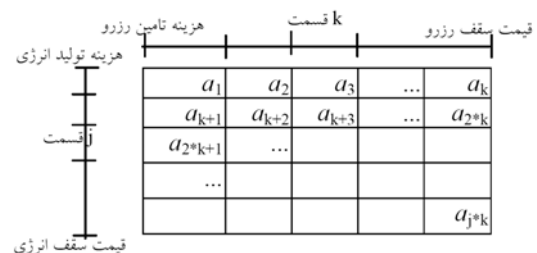
سیستم نمونه مورد مطالعه شامل ۷ تولیدکننده می باشد که مشخصات آنها در جدول (۱) ارائه شده است. همچنین بار پیش بینی شده فصلی توسط ISO برای ۲۴ ساعت روزانه در شکل (۵) نمایش داده شده است. در ضمن، میزان ظرفیت رزرو مورد نیاز سیستم در ۲۴ ساعت مقدار ثابت 60MW (برابر ظرفیت بزرگترین واحد موجود در سیستم) خواهد بود.

قیمت سقف بازار انرژی و رزرو چرخشی تقسیم خواهد شد. اتخاذ عمل  $a$ ، به معنای انتخاب یک قیمت تصادفی در بازه مورد نظر در فضای عمل خواهد بود.

شکل های (۳) و (۴)، به ترتیب شمای کلی از فضای حالت و فضای عمل در هر ساعت را نمایش می دهند.



شکل ۳: شمای کلی از فضای حالت ها در هر ساعت



شکل ۴: شمای کلی از فضای عمل ها برای هر عامل در هر ساعت

پاداش هر عامل در ساعت  $h$  ام تحت عمل  $a$  و حالت  $s$  به صورت زیر قابل محاسبه خواهد بود: [۱۶]

$$r_h(s, a) = r(h) \times \left( \frac{AUR(h)}{TUR(h)} \right)^b \quad (7)$$

که  $r(h)$  برابر مبلغ دریافتی عامل در ساعت  $h$  ام منهای هزینه های تامین در آن ساعت می باشد.  $AUR(h)$  و  $TUR(h)$  به ترتیب نرخ بکارگیری ظرفیت<sup>۱</sup> مطلوب و نرخ بکارگیری ظرفیت واقعی در ساعت  $h$  ام می باشند. پارامتر  $b$  نیز بیانگر اهمیت نرخ بکارگیری ظرفیت خواهد بود.

مراحل یادگیری و قیمت دهی هر عامل به شرح ذیل می باشد:

- **مرحله اول (تعیین حالت):** در شروع هر روز مبادلات برق، هر عامل بر اساس قیمت بسته شدن بازار انرژی و رزرو

1. Utilization rate

## بیست و پنجمین کنفرانس بین‌المللی برق

جدول ۲: پارامترهای تنظیمی عامل‌ها (واحدهای تولیدی)

Gen. No	$\alpha$	$\varepsilon$	$\gamma$	$b$	TUR
1	0.7	0.3	0.1	2	0.9
2	0.7	0.3	0.1	2	0.9
3	0.7	0.3	0.1	2	0.8
4	0.7	0.3	0.1	2	0.8
5	0.1	0.3	0.5	1	0.7
6	0.1	0.3	0.5	1	0.7
7	0.1	0.3	0.5	1	0.9

بر اساس جدول (۲)، نرخ بکارگیری ظرفیت مطلوب (TUR) هر عامل برای تمامی ساعت‌های شبانه روز یکسان فرض شده است.

## ۴-۲- مرحله یادگیری اولیه:

برای بدست آوردن مقادیر اولیه مناسب برای جدول ارزش حالت-عمل (Q)، ابتدا الگوریتم را برای ۱۰۰۰۰ روز مبادلات تحت هر دو مدل پرداخت (A) و (A+L) به طور جداگانه اجرا می‌نماییم. لازم به ذکر است، در طی پروسه یادگیری، مقدار ضریب  $\alpha$  برای هر عامل در هر ساعت و روز و در هر حالت و عمل به صورت زیر محاسبه می‌شود: [۱۶، ۲۲]

$$\alpha_{h,d}(s,a) = \frac{1}{\beta_{h,d}(s,a)} \quad (۸)$$

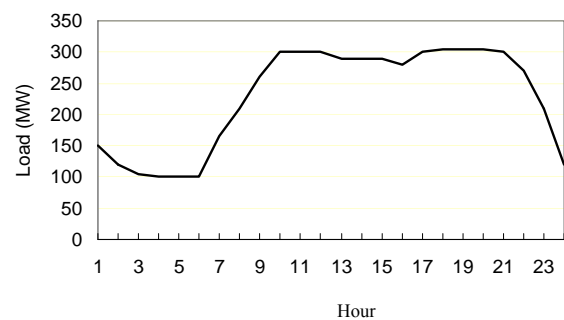
در رابطه فوق  $\beta_{h,d}(s,a)$  برابر تعداد دفعاتی است که به جفت حالت-عمل (s,a) تا روز مبادله مورد نظر رجوع شده است. البته باید توجه نمود که پس از یادگیری اولیه و در مرحله قیمت‌دهی اصلی، همان مقادیر ثابت  $\alpha$  که قبلاً در جدول (۲) برای عامل‌ها ذکر شد، مورد استفاده قرار خواهند گرفت.

## ۴-۳- مرحله قیمت‌دهی اصلی:

در این مرحله مقادیر اولیه ارزش حالت-عمل (Q) که در پروسه یادگیری بدست آمدند، مورد استفاده قرار می‌گیرند. الگوریتم قیمت‌دهی عامل‌ها در محیط بازار نیز تحت دو مدل پرداخت (A) و (A+L) به طور جداگانه برای ۲۰۰۰ روز اجرا می‌شوند.

جدول ۱: مشخصات واحدهای تولیدی سیستم مورد مطالعه

Gen. No	$P^{min}$ (MW)	$P^{max}$ (MW)	$SR^{max}$ (MW)	هزینه تولید انرژی (\$/MWh)	هزینه تامین ظرفیت رزرو چرخشی (\$/MW)
1	15	60	40	38	2
2	15	50	40	38	7
3	15	60	40	51	4.5
4	15	50	40	60	2
5	15	60	45	60	7
6	15	60	45	38	2
7	15	50	40	38	2



شکل ۵: بار پیش‌بینی شده فصلی توسط ISO برای ۲۴ ساعت

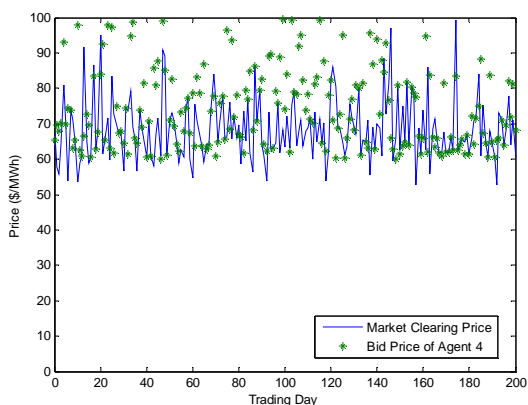
کف قیمت انرژی و رزرو چرخشی برای سیستم مورد مطالعه برابر صفر در نظر گرفته شده است. سقف قیمت در بازار انرژی و رزرو چرخشی نیز به ترتیب برابر 100\$/MWh و 50\$/MW خواهد بود.

قیمت بازار انرژی و رزرو چرخشی برای تعیین حالت به ترتیب به ۱۰ و ۵ بازه بین کف و سقف قیمت تقسیم شده‌اند. همچنین برای قیمت‌دهی عامل در بازار انرژی و رزرو چرخشی به ترتیب به ۱۰ و ۵ بازه بین هزینه مربوطه و سقف قیمت در نظر گرفته شده است. به عبارت دیگر با توجه به شکل‌های (۳) و (۴)،  $m=5$ ،  $n=10$ ،  $k=5$  و  $j=10$  خواهد بود.

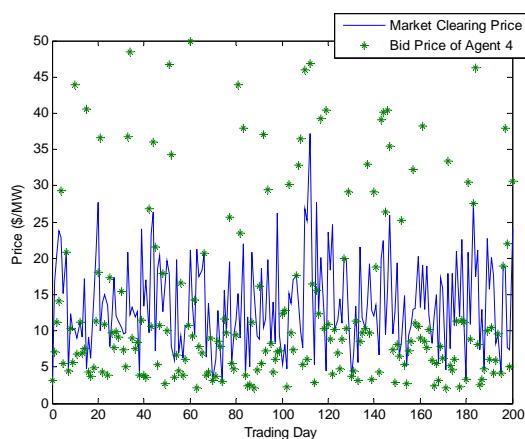
برخی پارامترهای تنظیمی برای هر یک از عامل‌های شرکت‌کننده در بازار نیز در جدول (۲) ارائه شده است. لازم به ذکر است، مقادیر پارامترها به جز  $\alpha$ ، هم در مرحله یادگیری اولیه و هم مرحله قیمت‌دهی اصلی عامل‌ها به کار گرفته خواهند شد. مقدار  $\alpha$  ارائه شده در جدول (۲)، تنها در مرحله قیمت‌دهی اصلی عامل‌ها استفاده خواهد شد و در مرحله یادگیری به طریق دیگری عمل می‌شود.



### بیست و پنجمین کنفرانس بین‌المللی برق



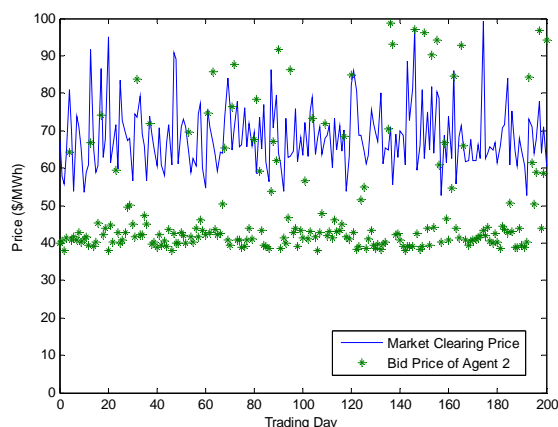
شکل ۸: قیمت بسته شدن بازار انرژی و همچنین قیمت پیشنهادی تولید کننده شماره ۴ در این بازار در ساعت ۷ ام تحت مدل (A)



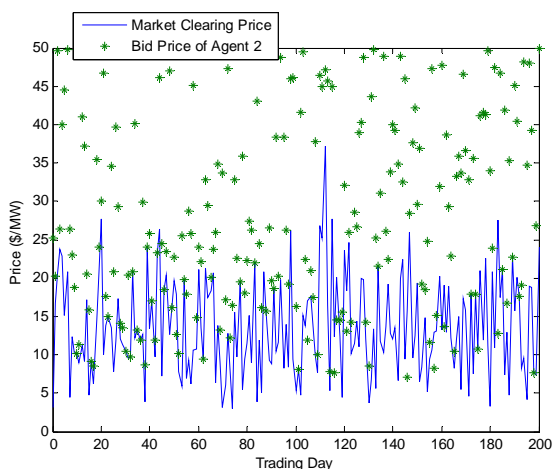
شکل ۹: قیمت بسته شدن بازار رزرو و همچنین قیمت پیشنهادی تولید کننده شماره ۴ در این بازار در ساعت ۷ ام تحت مدل (A)

با مقایسه شکل‌های (۶) و (۷) می‌توان دریافت، تولید کننده ۲ که نسبت به اکثر رقبا دارای هزینه تولید انرژی پایین‌تری است ولی هزینه تامین ظرفیت رزرو آن نسبتاً بالاست، بیشتر سعی خود را در برنده شدن در بازار انرژی معطوف کرده و قیمت پیشنهادی وی در این بازار در اکثر مواقع زیر قیمت بسته شدن بازار بوده است. این تولیدکننده که از پایین بودن قیمت خود در بازار انرژی و حصول سود کافی ناشی از دریافت قیمت بسته شدن این بازار اطمینان دارد، ریسک بیشتری را در بازار رزرو چرخشی متحمل شده و قیمت‌های پیشنهادی وی در این بازار در خیلی از مواقع برابر و یا بیشتر از قیمت بسته شدن این بازار می‌باشد. به عبارت دیگر این

جهت بررسی استراتژی‌های قیمت‌دهی، تولیدکننده‌های ۲ و ۴ به عنوان دو تولید کننده که هزینه تولید انرژی و هزینه تامین رزرو متفاوتی دارند، انتخاب شده‌اند. شکل‌های (۶) و (۷)، قیمت بسته شدن بازار انرژی و رزرو چرخشی و همچنین قیمت پیشنهادی تولید کننده شماره ۲ در این دو بازار را در ساعت ۷ ام تحت مدل پرداخت (A) نمایش می‌دهند. شکل‌های (۸) و (۹) نیز قیمت بسته شدن بازار انرژی و رزرو چرخشی و همچنین قیمت پیشنهادی تولید کننده شماره ۴ در این دو بازار را در ساعت ۷ ام تحت مدل پرداخت (A) نشان می‌دهند.



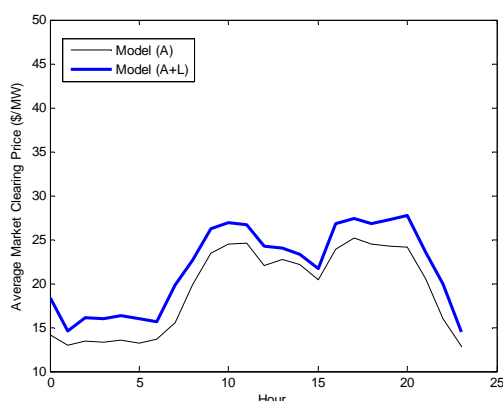
شکل ۶: قیمت بسته شدن بازار انرژی و همچنین قیمت پیشنهادی تولید کننده شماره ۲ در این بازار در ساعت ۷ ام تحت مدل (A)



شکل ۷: قیمت بسته شدن بازار رزرو و همچنین قیمت پیشنهادی تولید کننده شماره ۲ در این بازار در ساعت ۷ ام تحت مدل (A)



## بیست و پنجمین کنفرانس بین‌المللی برق



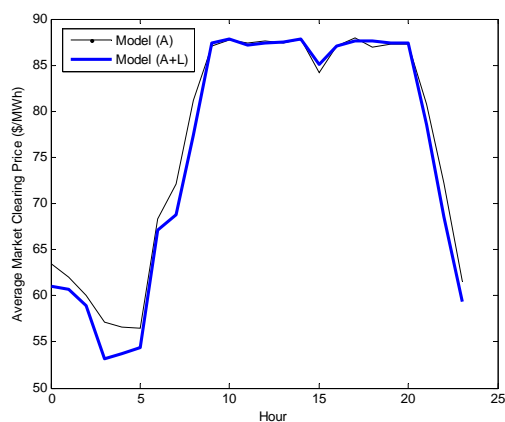
شکل ۱۱: متوسط قیمت بسته شدن بازار رزرو چرخشی تحت مدل (A) و مدل (A+L) برای ۲۴ ساعت

همانطور که در شکل (۱۰) مشخص است، قیمت بسته شدن بازار انرژی در حالت لحاظ نمودن هزینه فرصت از دست رفته (مدل (A+L)) در اکثر اوقات کمتر از حالتی می‌باشد که این هزینه در نظر گرفته نشده است (مدل (A)). این کاهش به طور متوسط در طول شبانه روز برابر 1.2 \$/MWh بوده است.

در مقابل، در شکل (۱۱) مشاهده می‌شود، قیمت بسته شدن بازار رزرو چرخشی در حالت لحاظ نمودن هزینه فرصت از دست رفته (مدل (A+L)) بیشتر از حالتی می‌باشد که این هزینه در نظر گرفته نشده است (مدل (A)). این افزایش به طور متوسط در طول شبانه روز برابر 2.6 \$/MWh بوده است. باید توجه کرد، در بازار مورد مطالعه که قانون تسویه آن یکنواخت است، پرداختی به برندگان مناقصات بازار بر اساس بالاترین قیمت پذیرفته شده محاسبه می‌شود. از طرف دیگر، میزان تقاضا در بازار انرژی معمولاً بیشتر از بازار رزرو است. لذا در بازاری با قانون تسویه یکنواخت، پایین‌تر بودن قیمت بسته شدن بازار انرژی به دلیل زیادتر بودن میزان تقاضا در این بازار، می‌تواند دارای اهمیت بیشتری باشد و میزان کل پرداختی به برندگان بازارها (هزینه کل تامین انرژی و رزرو) را کاهش دهد. بنابراین در نتایج حاصل از شبیه‌سازی، کاهش قیمت مشاهده شده در بازار انرژی تحت مدل (A+L) با وجود افزایش قیمت در بازار رزرو چرخشی، وضعیت بهتری را نوید می‌دهد. در واقع به نظر می‌رسد، در بازاری با قانون

عامل دریافته است که بازار انرژی برای وی اهمیت بیشتری دارد و شانس برنده شدن بیشتری در آن خواهد داشت. همچنین با مقایسه شکل‌های (۸) و (۹) می‌توان دریافت، تولیدکننده ۴ که نسبت به اکثر رقبا دارای هزینه تامین رزرو پایین‌تری است ولی هزینه تولید انرژی آن نسبتاً بالاست، بیشتر سعی خود را در برنده شدن در بازار رزرو چرخشی معطوف کرده و قیمت پیشنهادی وی در این بازار در اکثر مواقع زیر قیمت بسته شدن بازار بوده است. قیمت‌های پیشنهادی این تولیدکننده در بازار انرژی، در خیلی از مواقع برابر قیمت بسته شدن این بازار می‌باشد و به عبارت دیگر این تولیدکننده در بسیاری از روزها نقش تعیین‌کننده‌ای در قیمت بسته شدن بازار انرژی ایفا کرده است. در واقع این عامل دریافته است که بازار رزرو چرخشی برای وی اهمیت بیشتری دارد ولی برای برنده شدن در این بازار حتماً باید در بازار انرژی نیز برنده باشد.

پس از ۲۰۰۰ تکرار، یک متوسط گیری برای ۲۴ ساعت انجام شده و قیمت بازار انرژی و رزرو چرخشی برای ۲۴ ساعت در دو مدل پرداخت (A) و (A+L) رسم شدند. شکل‌های (۱۰) و (۱۱) به ترتیب متوسط قیمت بسته شدن بازار انرژی و بازار رزرو چرخشی را در دو مدل مورد بحث نمایش می‌دهند.



شکل ۱۰: متوسط قیمت بسته شدن بازار انرژی تحت مدل (A) و مدل (A+L) برای ۲۴ ساعت

## بیست و پنجمین کنفرانس بین‌المللی برق

## ۵- جمع بندی و نتیجه‌گیری:

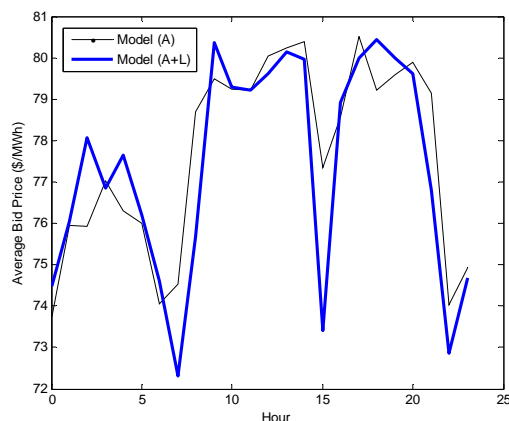
در این مقاله، به کمک الگوریتم Q-Learning روشی مبتنی بر سیستم چند عاملی جهت شبیه‌سازی مناقصه همزمان بازار انرژی و رزرو چرخشی پیشنهاد شد که از طریق آن دو مدل پرداخت با هزینه فرصت از دست رفته و بدون آن، در خصوص یک سیستم نمونه مورد تجزیه و تحلیل و مقایسه قرار گرفتند. لازم به ذکر است، اجرای مناقصات بازارهای انرژی و رزرو در سیستم مورد مطالعه به صورت Day-Ahead و بر اساس قانون تسویه یکنواخت در نظر گرفته شدند. نتایج شبیه‌سازی بر روی سیستم نمونه نشان می‌دهند، لحاظ هزینه فرصت از دست رفته در مدل پرداختی به برندگان می‌تواند در سیستمی با قانون تسویه یکنواخت، اثرات مثبتی از قبیل کاهش قیمت بسته شدن بازار انرژی و متعاقباً کاهش میزان کل پرداختی به برندگان بازارهای انرژی و رزرو را در پی داشته باشد.

## مراجع:

- [1] Oren, S.S., "Design of Ancillary Service Markets", Proceedings of the 34<sup>th</sup> Hawaii International Conference on System Sciences, 2001.
- [2] Xu, Z., Dong, Z.Y. and Wong, K.P., "Optimal Dispatch of Spinning Reserve in a Competitive Electricity Market Using Genetic Algorithm", Evolutionary Computation, IEEE, Vol. 1, pp. 587-592, December 2003.
- [3] Verbic, G. and Gubina, F., "Cost-Based Models for the Power-Reserve Pricing of Frequency control", IEEE Transactions on Power Systems, Vol. 19, No. 4, pp. 1853-1858, November 2004.
- [4] Rashidi-Nejad, M., Song, Y.H. and Javidi-Dasht-Bayaz, M.H., "Operating Reserve Provision in Deregulated Power Markets", Power Engineering Society Winter Meeting, IEEE, Vol. 2, pp. 1305-1310, January 2002.
- [5] Prada, R.B., et al, "Possible Design of The Brazilian Ancillary Service Market for Generation Reserves", Power System Management and Control, Conference Publication No. 488, pp. 414-419, April 2002.
- [6] Allen, E.H. and Ilic, M.D., "Reserve Market for Power Systems Reliability", IEEE Transactions on Power Systems, Vol. 15, No. 1, pp. 228-233, February 2000.
- [7] Chen, J., et al, "Locational Pricing and Scheduling for an Integrated Energy-Reserve Market", Proceedings of the 36th Hawaii International Conference on System Science (HICSS'03), 2002.

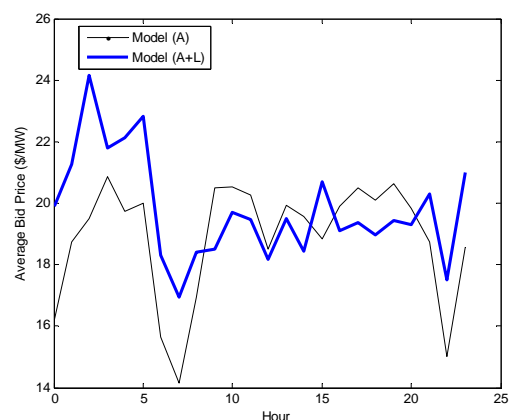
تسویه یکنواخت، لحاظ نمودن هزینه فرصت از دست رفته در مدل پرداخت، اثرات مثبتی از جمله کاهش هزینه تامین کل را به دنبال داشته باشد.

لازم به ذکر است، علاوه بر اینکه مسئله بهینه‌سازی که ISO حل می‌کند در دو مدل پرداخت مورد بحث متفاوت بوده است، استراتژی قیمت‌دهی تولیدکنندگان نیز در این دو مدل اختلاف‌هایی دارد. شکل‌های (۱۲) و (۱۳) که متوسط قیمت‌دهی یک تولیدکننده نمونه (شماره ۴) به ترتیب در بازار انرژی و رزرو را تحت مدل (A) و (A+L) نمایش می‌دهند، مبین این واقعیت هستند.



شکل ۱۲: متوسط قیمت پیشنهادی تولیدکننده شماره ۴ در بازار انرژی

تحت مدل (A) و مدل (A+L) برای ۲۴ ساعت



شکل ۱۳: متوسط قیمت پیشنهادی تولیدکننده شماره ۴ در بازار رزرو

چرخشی تحت مدل (A) و مدل (A+L) برای ۲۴ ساعت

بیست و پنجمین کنفرانس بین‌المللی برق

- [21] Rahimiyan, M. and Rajabi Mashhadi, H., "Evaluating the efficiency of divestiture policy in promoting competitiveness using an analytical method and agent-based computational economics", Energy Policy, Vol. 38, Issue 3, pp. 1588-1595, March 2010.
- [22] Nie, J. and Haykin, S., "A dynamic Channel Assignment Policy through Q-Learning", IEEE Trans. Neural Networks, vol. 10, No. 6, pp. 1443-1455, Nov. 1999.
- [8] Zhu, J., et al, "The Market for Spinning Reserve and Its Impacts on Energy Prices", Power Engineering Society Winter Meeting, IEEE, Vol. 2, pp. 1202 - 1207, January 2000.
- [9] Liao, Y., Feng, X. and Pan, J., "Analysis of Interaction between Ancillary Services Markets and Energy Market Using Power Market Simulator", IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT), pp. 117-122, April 2004.
- [10] Market Design Department, "Ancillary Services Market Enhancements: Reserves White Paper", ISO New England Inc, May 2004.
- [11] FERC Electric Tariff, "Order on Further Development of The California ISO's Market Redesign and Establishing Hearing Procedures", California Independent System Operator Corporation, Docket No. ER02-1656 - 017, June 2004.
- [12] Liu, Y., et al, "Implementing Rational Buyer's Algorithm at California ISO", Power Engineering Society Winter Meeting, IEEE, Vol. 2, pp. 1293-1298, January 2002.
- [13] Gan, D. and Litvinov, E., "Energy and Reserve Market Designs With Explicit Consideration to Lost Opportunity Costs", IEEE Transactions on Power Systems, Vol. 18, No. 1, pp. 53-59, February 2003.
- [14] Chen, J., et al, "Coordinated Interchange Scheduling and Opportunity Cost Payment: A Market Proposal to Seams Issues", Proceedings of the 37<sup>th</sup> Hawaii International Conference on System Science, 2004.
- [15] Sutton, R.S. and Barto, A.G., "Reinforcement learning: an introduction", The MIT Press, A Bradford Book, 1998.
- [16] Xiong, G., Okuma, S. and Fujita, H., "Multi-agent Based Experiments on Uniform Price and Pay-as-Bid Electricity Auction Market", IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT'2004), Vol. 1, pp. 72-76, 2004.
- [17] Xiong, G., Hashiyama, T. and Okuma, S., "A Supplier Bidding Strategy Through Q-Learning Algorithm in Electricity Auction Markets", IEEE Trans. Electronics, Information and Systems, vol. 123, No. 6, pp. 1134-1141, June 2003.
- [18] Tellidou, A.C. and Bakirtzis, A.G., "Multi-Agent Reinforcement Learning for Strategic Bidding in Power Markets", 3rd International IEEE Conference Intelligent Systems, pp. 408-413, September 2006.
- [19] Rahimiyan, M. and Rajabi Mashhadi, H., "Supplier's optimal bidding strategy in electricity pay-as-bid auction: Comparison of the Q-learning and a model-based approach", Electric Power Systems Research 78 (2008), pp. 165-175, 2008.
- [20] Rahimiyan, M. and Rajabi Mashhadi, H., "An Adaptive Q-Learning Algorithm Developed for Agent-Based Computational Modeling of Electricity Market", IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, Issue 99, pp. 1-10, 2010.