

## KẾT HỢP THUẬT TOÁN SA VÀ Q-LEARNING ÁP DỤNG TRONG THỊ TRƯỜNG ĐIỆN BÁN BUÔN CẠNH TRANH

BẠCH THANH QUÝ, PHAN LÂM VŨ, VĂN THỊ KIỀU NHI  
Khoa Công nghệ Điện, Đại học Công nghiệp Thành phố Hồ Chí Minh  
bachthanhquy@iuh.edu.vn

**Tóm tắt:** Thị trường điện Việt Nam đã chuyển sang cấp độ 2 – Cấp độ thị trường điện bán buôn cạnh tranh. Các nhà máy điện (Gencos) và các công ty kinh doanh phụ tải điện (LSEs) tham gia thị trường bán buôn cạnh tranh với mục tiêu tối đa hóa lợi nhuận. Chiến thuật chào giá tối ưu được các đơn vị tham gia được triệt để khai thác. Bài viết này trình bày thị trường điện bán buôn cạnh tranh dạng mô hình động. Trong môi trường cạnh tranh thông tin không hoàn hảo, các thuật toán tối ưu được sử dụng để tối ưu hóa bản chào giá. Thuật toán mô phỏng luyện kim (SA - Simulated annealing algorithm) kết hợp với thuật toán học Q tăng cường (Q – Learning reinforcement algorithm) thành thuật toán SA – Q learning được nhóm tác giả đề xuất áp dụng cho chiến lược chào giá tối ưu cho các đơn vị cung cấp điện. Sơ đồ tiêu chuẩn IEEE – 30 nút được sử dụng để mô phỏng mức độ đáp ứng của thuật toán đề xuất.

**Từ khóa:** Thị trường điện, thuật toán Q-learning, thuật toán SA, thị trường điện bán buôn.

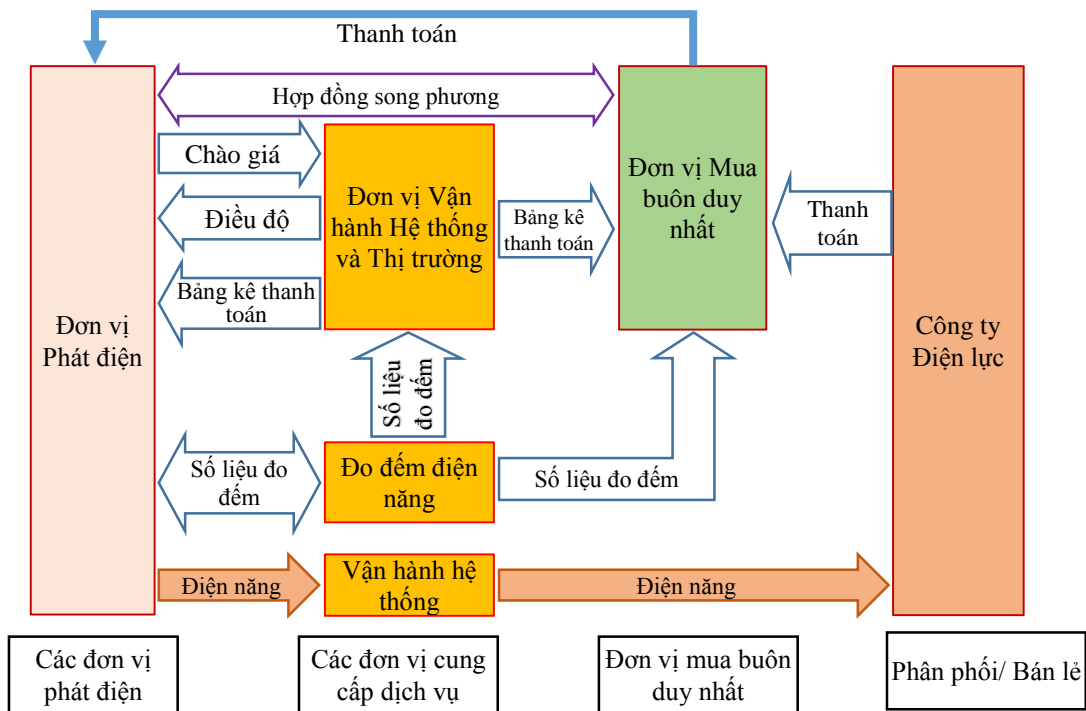
### COMBINED THE SIMULATED ANNEALING AND THE Q-LEARNING ALGORITHM APPLICATION TO THE WHOLESALE ELCTRICITY MARKET

**Abstract:** Vietnam's electricity market has moved to the level 2 – The competitive wholesale electricity market level. The generation companies (GenCos) and the load service entities (LSEs) participate in the market with aim maximum profits. The optimal bidding strategy is a key for exploitation by the participants. In this article, The dynamic competitive wholesale power market model is presented. The optimization algorithms are used to solve the optimal bidding strategy problems. The SA-Q learning algorithm is a combined of the simulated annealing algorithm and the Q – reinforcement learning algorithm, that is proposed and used for GenCos to bidding strategy. The IEEE 30-nút test system with six Gencos and three LSEs is used for simulations.

**Keywords:** Electricity market, Q-learning algorithm, SA algorithm, Wholesale electricity market

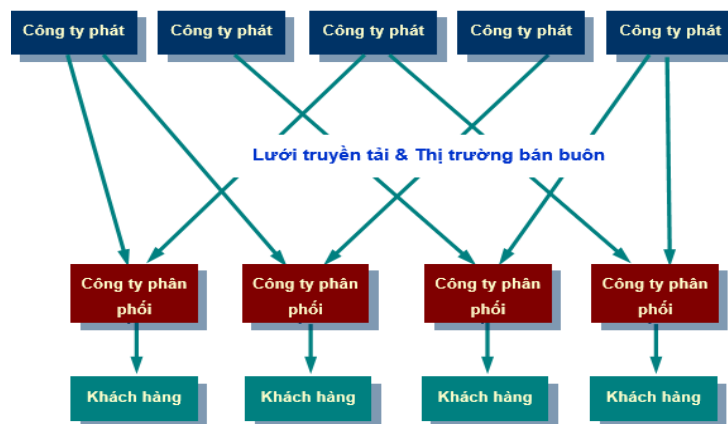
#### 1 ĐẶT VẤN ĐỀ

Trong [1], Thị trường điện Việt Nam được phê duyệt lộ trình phát triển theo 3 cấp độ: Cấp độ 1 là thị trường phát điện cạnh tranh (VCGM – Vietnam Competition Generation Market), cấp độ 2 là thị trường bán buôn cạnh tranh (VWEM – Vietnam Wholesale Electricity Market), và cấp độ 3 là thị trường bán lẻ cạnh tranh (VDEM - Vietnam Detail Electricity Market). VCGM chính thức đi vào hoạt động từ ngày 01/7/2012 theo mô hình thị trường chào giá tập trung theo chi phí biến đổi. Điện năng của các nhà máy điện được bán cho đơn vị mua buôn duy nhất đó chính là Công ty Mua bán điện thuộc Tập đoàn Điện lực Việt Nam. Mô hình VCGM được trình bày như Hình 1 sau:



Hình 1: Cấu trúc thị trường VCGM

Theo [2], thị trường điện Việt Nam đã chuyển sang cấp độ thị trường cấp độ 2 – cấp độ bán buôn điện cạnh tranh. VWEM sau giai đoạn tính toán mô phỏng và vận hành thí điểm đã chính thức chuyển sang thanh toán thật từ 1/2019. Mô hình bán buôn điện cạnh tranh được trình bày như Hình 2.



Hình 2: Mô hình thị trường VWEM

Trong VWEM không có công ty mua buôn duy nhất như mô hình VCGM mà có nhiều đơn vị mua buôn tham gia cạnh tranh với nhau. Đơn vị vận hành thị trường (Market operator - MO) và đơn vị vận hành hệ thống độc lập (Independent system operator - ISO) là hai đơn vị đóng vai trò trọng tài của cuộc chơi giữa một bên là các công ty phát điện (Generation companies - GenCos) và một bên là các đơn vị mua buôn phụ tải điện (Load service entities - LSEs). MO và ISO không tham gia kinh doanh hay bán buôn điện, chỉ đảm nhiệm minh bạch thị trường và vận hành hệ thống điện để đảm bảo an toàn, an ninh năng lượng cho hệ thống điện. Những đơn vị tham gia thị trường (gồm GenCos và LSEs) sẽ có điều kiện cạnh tranh giá bán buôn với nhau, và tìm kiếm cơ hội gia tăng lợi nhuận. Hai bên đại diện cho hai đối trọng của cân bằng kinh tế đó là cung và cầu, giá điện được hình thành dựa trên cân bằng cung cầu này.

Trong môi trường thị trường điện cạnh tranh, các nhà máy điện muốn nâng cao lợi nhuận của mình cần phải có chiến lược chào giá tối ưu. Vì thị trường giá điện được quyết định dựa trên cơ sở chào giá bán của các nhà máy điện và chào giá mua của các đơn vị mua buôn. Do đó hành vi chào giá của các đơn vị không chỉ ảnh hưởng trực tiếp đến khả năng được huy động phát của nhà máy mà còn có thể tác động đến giá của toàn hệ thống. Ý thức được việc ảnh hưởng này, nên thông tin của các đối thủ trong hệ thống hoàn toàn

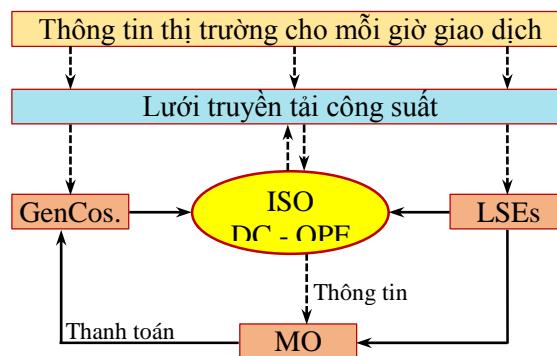
không được biết. Việc thiếu thông tin là một trở ngại cho việc quyết định phương án chào giá. Nhiều thuật toán được đề xuất cho bài toán tối ưu trong môi trường thông tin không hoàn hảo.

Nâng cao khả năng cạnh tranh trong chiến lược chào giá tối ưu cho các đơn vị tham gia thị trường được nghiên cứu nhiều trong những năm gần đây. Lý thuyết trò chơi áp dụng trong chiến lược chào giá tối ưu cũng đã được nghiên cứu. Trong [3] và [4], lý thuyết tiếp cận trò chơi Nash được áp dụng cho chiến thuật chào giá tối ưu trong thị trường điện điều tiết, nơi mà các đơn vị tham gia thị trường thiếu thông tin đối thủ. Tuy nhiên, kết quả mô phỏng cho thấy chiến lược chào giá tối ưu thay đổi theo cấp độ thông tin mà đơn vị tham gia có được về đối thủ cạnh tranh. Lý thuyết trò chơi bất hợp tác của Cournot cũng đã được đưa ra áp dụng để xác định lượng công suất phát tối ưu cho các nhà máy điện tham gia trong thị trường điện độc quyền được nêu ra trong [5]. Kết quả cho thấy ước tính độ chính xác hàm chi phí sản xuất của đối thủ đóng vai trò quyết định trong thị trường này. Quá trình quyết định Markov cho chiến lược chào giá tối ưu trong thị trường điện giao ngay được đề xuất trong [6]. Thuật toán di truyền được phát triển trong chiến lược chào giá thầu cho những đơn vị tham gia trong thị trường đấu giá hai chiều cũng được trình bày trong [7] và [8]. Thuật toán mờ và mạng nơ ron nhân tạo áp dụng cho chiến lược chào giá tối ưu trong thị trường điện cạnh tranh cũng được phát triển trong [9]. Thuật toán học Q- learning cũng được sử dụng để giải quyết chiến lược chào giá tối ưu cho những đơn vị tham gia thị trường [10] và [11]. Thuật toán học Q – learning cũng đã được áp dụng cho các máy điện thực hiện chiến lược chào giá tối ưu khi tham gia cạnh tranh trong thị trường cạnh tranh thông tin không hoàn chỉnh được đề xuất trong [12],[13], và [14]. Kết quả về mức độ hội tụ, phù hợp của thuật toán trong thị trường điện cạnh tranh cũng đã được phân tích và chứng minh.

Trong bài viết này, mô hình thị trường VWEM trước ngày được trình bày, xác định chức năng các đơn vị tham gia và trình bày cơ chế vận hành thị trường. Giá điện bán buôn được xác định dựa trên quy luật cân bằng cung cầu thị trường, khi giao dịch đã xác lập thì giá biên được công bố để làm minh bạch thông tin, tạo cơ chế “cuộc chơi” cạnh tranh lành mạnh. Đơn vị ISO không tham gia cạnh tranh, chỉ đảm trách vai trò đảm bảo kỹ thuật vận hành tối ưu hoá công suất trên hệ thống bằng phương pháp tối ưu trào lưu công suất DC – OPF (Direct current – Optimal power flow) để xác định công suất giao dịch tại các nút. Trong bài viết này, để tối ưu hoá lợi nhuận của các đơn vị tham gia cạnh tranh, nhóm tác giả đề xuất sử dụng thuật toán mô phỏng luyện kim (SA - Simulated annealing algorithm) kết hợp với thuật toán học Q tăng cường (Q – Learning reinforcement algorithm) thành thuật toán SA – Q learning áp dụng tìm chiến lược chào giá tối ưu. Sơ đồ tiêu chuẩn IEEE – 30 nút được sử dụng để mô phỏng mức độ đáp ứng của thuật toán.

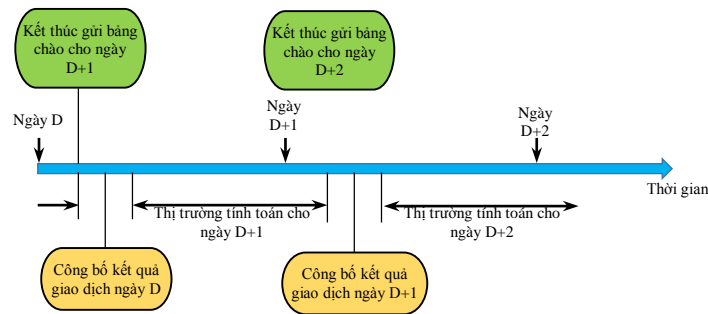
## 2 MÔ HÌNH THỊ TRƯỜNG ĐỘNG

Giá điện bán buôn được hình thành từ kết quả thay đổi liên tục của cấp độ cung và cầu trên thị trường, mô hình thị trường động được xây dựng chi tiết trên cơ sở của mô hình bán buôn cạnh tranh VWEM, không gian thị trường thay đổi không ngừng. Mô hình thị trường VWEM được triển khai chi tiết như Hình 3.



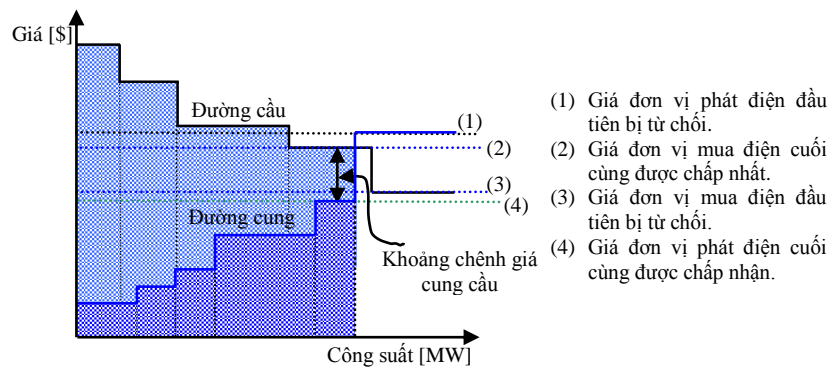
Hình 3: Cấu trúc hoạt động của VWEM

Trong hình 3, ISO và MO đóng vai trò đảm bảo giao dịch, giao dịch được thực hiện trong từng giờ trước 1 ngày, thông tin giao dịch thành công được công bố công khai. Trên cơ sở thông tin đã giao dịch trước đó, GenCos và LSEs tiến hành lựa chọn chiến lược tốt nhất để chào giá tham gia thị trường. ISO tính toán lượng công suất và giá tương ứng tại các nút trên hệ thống sau khi đã nhận được bản chào giá mua và giá chào bán của các bên tham gia. Giá giao dịch và công suất giao dịch trong từng giờ của những đơn vị giao dịch thành công được công bố. Toàn bộ giao dịch của ngày trước được thực hiện trong ngày hôm nay nên mô hình này được gọi tên là mô hình thị trường bán buôn trước ngày. Chi tiết được mô tả trong Hình 4.



Hình 4: Mô hình thị trường bán buôn trước ngày

Trong ngày D, tại giờ t(h), ISO sẽ nhận giá thầu và giá cung cấp của tất cả thành viên tham gia cho 24h giao dịch của ngày D+1, sau đó ISO sẽ tiến hành tính toán số lượng công suất và giá tương ứng tại các nút giao dịch trên hệ thống. ISO có nhiệm vụ đảm bảo cho hệ thống hoạt động tin cậy, ổn định và công bằng. Có nhiều giải thuật trong phân bố công suất, tuy nhiên trong mô hình này tác giả chọn giải pháp DC-OPF để tính toán bài toán phân bố công suất cân bằng trên hệ thống. Kết quả giao dịch trong ngày D+1 được công bố đầu ngày D+1. Bắt đầu của ngày D+1, các giao dịch thành công của ngày D+1 phải cam kết thực hiện để đảm bảo ổn định hệ thống. Bên cạnh đó bộ phận tham gia thị trường cũng chuẩn bị giao dịch cho ngày kế tiếp D+2. Để đồng nhất giá biên trên thị trường, mô hình thị trường giao dịch hai bên được đề xuất như hình 5.



Hình 5: Giá biên trên thị trường giao dịch hai bên

Giá chào của các Gencos và giá thầu của các LSEs là giá không co giãn, giá chào được sắp xếp theo thứ tự tăng dần và giá thầu được sắp xếp theo thứ tự giảm dần. Với một số thị trường áp dụng, đơn vị vận hành thị trường có thể chọn một trong bốn căn cứ như Hình 5 làm giá giao dịch cuối cùng đó là: Giá đơn vị phát điện đầu tiên bị từ chối; giá đơn vị mua điện cuối cùng được chấp nhận; giá đơn vị mua điện đầu tiên bị từ chối; và giá đơn vị phát điện cuối cùng được chấp nhận. Trong bài viết này, giá biên giao dịch sẽ do MO quyết định và nằm trong khoảng chênh lệch giá cung cầu.

### 2.1 Giá chào

Mục tiêu của các Gencos trong thị trường cạnh tranh là tối đa hoá lợi nhuận, điều kiện trước chào giá là môi trường thiếu thông tin về đối thủ. Giá chào của các Gencos được xác định:

$$C_i(P_{gi}) = a_i P_{gi}^2 + b_i P_{gi} + c_i \quad (1)$$

$$C_{mi} = \frac{\partial C_i}{\partial P_{gi}} = 2a_i P_{gi} + b_i \quad (2)$$

$$P_{gi\_min} \leq P_{gi} \leq P_{gi\_max} \quad (3)$$

Trong đó  $a_i$ ,  $b_i$ ,  $c_i$  là các hệ số của hàm giá thực  $C_i(P_{gi})$  của máy phát thứ  $i$ , phát ra lượng công suất tác dụng  $P_{gi}$ .  $P_{gi\_min}$  và  $P_{gi\_max}$  là giới hạn trên và giới hạn dưới của máy phát thứ  $i$ , công thức (2) là đạo hàm giá theo công suất của máy phát được gọi là chi phí cận biên thực của máy phát thứ  $i$ .

Với mỗi giao dịch, Gencos cung cấp đến ISO và MO bảng chào chứa đựng hai thành phần đó là lượng công suất phát và giá tương ứng. Các bảng chào này được tính toán từ giá phát theo công thức (1) đến giới hạn công suất phát theo công thức (3). Hàm giá chào được xác định như sau:

$$f_i^{offer} = \{(q_i^1; A_i^k \otimes p_i^1), \dots, (q_i^n; A_i^k \otimes p_i^n)\} \quad (4)$$

$$A_{i\_min}^k \leq A_i^k \leq A_{i\_max}^k \quad (5)$$

$$A_i^k = A_{i\_min}^k + \frac{k}{h} (A_{i\_max}^k - A_{i\_min}^k) \quad (6)$$

Trong đó  $q_i$ ,  $p_i$  là công suất và giá của máy phát thứ  $i$ .  $A_i^k$  là hành động thứ  $k$  trong tập hành động của máy phát thứ  $i$ . Dựa trên hành động  $A_i^k$  mà máy phát thứ  $i$  có thể thay đổi giá trong hàm chào giá  $f_i^{offer}$  (4), và  $h$  là số hành động có thể được lựa chọn ( $k = 1, 2, \dots, h$ ).

ISO và MO tính toán và công bố giá điện tại nút giao dịch  $N$  trong thời điểm  $t$  được xác định là  $\lambda_{p,t}^N$ . Giá trị lợi nhuận thu được của máy phát thứ  $i$  trong 24h giao dịch được xác định:

$$r_{i,t} = \sum_{t=1}^{24} (\lambda_{p,t}^N P_{gi,t} - C_i(P_{gi,t})) \quad (7)$$

Trong đó  $r_{i,t}$  là hàm số lợi nhuận của máy phát  $i$  trong giờ  $t$ .

## 2.2 Giá Thầu

Tương tự như Gencos, hàm lợi nhuận của các LSEs được xác định như sau:

$$B_j(P_{dj}) = e_j P_{dj} - f_j P_{dj}^2 \quad (8)$$

$$B_{mj} = \frac{\delta B_j}{\delta P_{dj}} = e_j - 2f_j P_{dj} \quad (9)$$

$$\sum P_{dj} = \sum P_{gi} \quad (10)$$

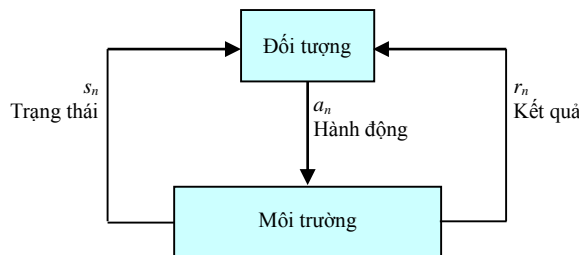
Trong đó  $e_j$ ,  $f_j$  là các hệ số của hàm lợi nhuận  $B_j(P_{dj})$  của LSEs thứ  $j$  và  $P_{dj}$  là công suất tác dụng mua được từ thị trường của LSEs thứ  $j$ .

Mỗi giao dịch LSEs cung cấp cho ISO và MO một bản chào giá thầu, trong đó có chứa hai thành phần đó là lượng công suất và giá tương ứng có nhu cầu mua. Bảng chào giá thầu được tính toán từ hàm lợi nhuận (8) và lượng công suất tải cần tiêu thụ. Hàm giá thầu được xác định như sau:

$$f_j^{bid} = \{(q_j^1; p_j^1), \dots, (q_j^n; p_j^n)\} \quad (11)$$

## 3 THUẬT TOÁN HỌC Q - LEARNING

Thuật toán học tăng cường Q-learning được đề xuất bởi Watkins để giải quyết quá trình quyết định Markov trong môi trường thông tin không đầy đủ. Ý tưởng chủ đạo của thuật toán học là học cách sinh tồn của loài vật trong môi trường tự nhiên. Các đối tượng trong tự nhiên phải tương tác theo môi trường đang tồn tại. Đối tượng sẽ nhận được một kết quả sau mỗi hành động, và rút ra kinh nghiệm để tồn tại. Ưu điểm lớn nhất của ý tưởng thuật toán này là tính đơn giản, dễ áp dụng và trực tiếp lên đối tượng. Những đơn vị tham gia sử dụng thuật toán học Q – learning để thực hiện chiến lược chào giá tối ưu trong thị trường điện bán buôn cạnh tranh trước ngày trong môi trường thiếu thông tin đối thủ.



Hình 6: Mô hình thuật toán học tăng cường

Như trình bày ở Hình 6, các đối tượng nhận được một kết quả  $r_n$  tương ứng với một cặp hành động – trạng thái  $(a_n, s_n)$  ở hành động thứ  $n$ . Hàm giá trị của cặp hành động – trạng thái được lưu lại trong bảng giá trị kinh nghiệm  $Q$  và luôn được cập nhật trong từng hành động.

Gọi  $S = \{s_1, s_2, \dots, s_m\}$  là một tập chứa  $m$  giá trị trạng thái môi trường và  $A = \{a_1, a_2, \dots, a_k\}$  là một tập  $k$  hành động mà đối tượng có thể thực hiện. Giả sử ở bước thứ  $n$ , tương ứng với trạng thái môi trường  $s_n \in S$ , đối tượng chọn cho mình một hành động tốt nhất  $a_n \in A$ , thì ngay lập tức nhận được một kết quả  $r_n$ . Đối tượng tiếp tục xem xét trạng thái môi trường tiếp theo  $s_{n+1} \in S$  và cập nhật lại giá trị kinh nghiệm  $Q$  tương ứng:

$$Q_{n+1}(s, a) = \begin{cases} (1 - \alpha)Q_n(s, a) + \alpha[r_n + \gamma \cdot \max_a Q_n(s_{n+1}, a')] & \text{if } s = s_n \text{ and } a = a_n \\ Q_n(s, a) & \text{otherwise} \end{cases} \quad (12)$$

Trong (12),  $\alpha \in (0, 1]$  là hệ số học và  $\gamma \in [0, 1)$  là hệ số suy giảm. Ý nghĩa của hệ số suy giảm là giá trị nhận được ở tương lai gần lớn hơn giá trị nhận được trong tương lai xa. Thuật toán SA-Q sử dụng phương pháp  $\varepsilon$  – greedy được đề xuất theo các bước:

*Begin:*

1. Nhận diện tất cả các cặp giá trị  $Q(s, a)$ .
2. Vòng lặp (cho mỗi cặp giá trị):
  - a. Lựa chọn ngẫu nhiên hành động  $a_r \in A$ .
  - b. Lựa chọn hành động theo quy luật ( $\varepsilon$  – greedy)  $a_p \in A$ :

$$a_p = \arg \max_a Q_{n-1}(s_n, a) \quad (13)$$

- c. Tạo giá trị ngẫu nhiên  $\varepsilon \in (0, 1)$ .
- d. Lựa chọn và thực hiện hành động  $a_n \in A$  theo quy luật gần đúng ngẫu nhiên:

$$a_n = \begin{cases} a_p & \text{if } \varepsilon \geq \exp\left[\frac{Q_{n-1}(s_n, a_r) - Q_{n-1}(s_n, a_p)}{\text{Temperature}}\right] \\ a_r & \text{otherwise} \end{cases} \quad (14)$$

- e. Nhận giá trị phản hồi  $r$  và cập nhật lại trạng thái mới  $s'$
  - f. Cập nhật lại bảng giá trị  $Q$  theo công thức (12)
- Tính toán lại giá trị temperature bằng hệ số suy giảm Temperature

3. Cho đến khi đạt được giá trị mục tiêu.

*End.*

Gọi  $T_n$  là hệ số Temperature tại bước thứ  $n$ ,  $T_{n+1} = \theta \times T_n$  là hệ số Temperature tại bước kế tiếp  $n+1$ . Hệ số suy giảm  $\theta \in (0.5, 1)$  quyết định tốc độ hội tụ nhanh hay chậm của thuật toán.

#### 4 ÁP DỤNG SA-Q LEARNING CHO THỊ TRƯỜNG ĐIỆN

Giả định các đơn vị tham gia thị trường với mục tiêu tối đa hóa lợi nhuận và giảm thiểu rủi ro. Để đưa thuật toán SA-Q learning áp dụng vào các đơn vị tham gia, cần định nghĩa các tập trạng thái, hành động và kết quả.

**Trạng thái (s):** Trạng thái được xác định là các giá điện trong thị trường.

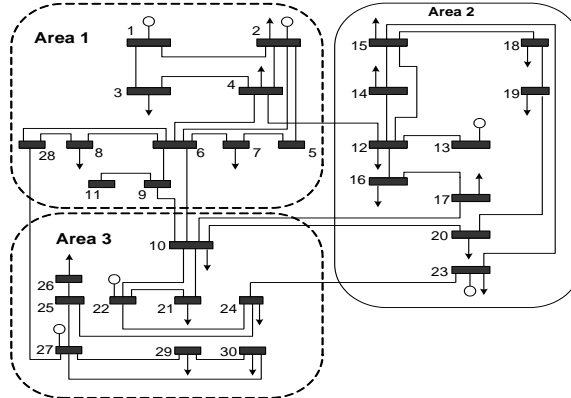
**Hành động (a):** Mỗi đơn vị tham gia có một tập hành động, chi tiết của hành động theo công thức (6) được xác định chi tiết:

$$A_{i\_min}^k = 1; A_{i\_max}^k = 2; h = 10$$

**Kết quả (r):** Sau khi ISO và MO tính toán và công bố công suất và giá tương ứng tại các nút trên hệ thống, mỗi đơn vị tham gia tính toán lợi ích thu được dựa trên hàm giá và lượng công suất cam kết phát được đưa lên hệ thống theo công thức (7). Theo [15], các thông số thuật toán ảnh hưởng không nhỏ đến tốc độ hội tụ của thuật toán, do vậy trong bài viết này, nhóm tác giả chọn không thay đổi thông số thuật toán đó là: Hệ số học  $\alpha = 0.5$ ; hệ số suy giảm  $\gamma = 0.1$ ; hệ số Temperature  $T = 100,000$ ;  $\theta = 0.99$ ; giá trị ban đầu của bảng học  $Q_0(s_0, a_0) = 0$ .

## 5 MÔ PHỎNG

Sử dụng hệ thống điện tiêu chuẩn IEEE 30 nút với 6 nhà máy điện thực hiện mô phỏng phân bố công suất và vận hành thị trường điện. Sơ đồ mạng điện như Hình 7



Hình 7: Sơ đồ đơn tuyến của lưới truyền tải IEEE 30 nút [16]

Thông tin cơ bản của hệ thống được trình bày như trong Bảng 1

Bảng 1: Thông tin cơ bản của hệ thống

Tổng công suất máy phát	352MW
Tải cố định	151.64MW
Tải có thể điều chỉnh	90MW

Phân bố công suất được tính toán theo thuật toán phân bố công suất tối ưu DC-OPF trong môi trường Matlab. 6 nhà máy điện trong hệ thống được phân bố trong 3 vùng, trên các nút số 01, 02, 13, 23, 22 và 27 được đề xuất sử dụng thuật toán SA-Q learning với cùng bộ thông số. Bảng 2 và bảng 3 là các bảng giá chào và giá thầu của 6 nhà máy điện và 3 đơn vị mua buôn tải.

Bảng 2: Thông tin bảng giá chào của 6 GenCos

Gen	Nút	$P_g$ Min	$P_g$ max	Block 1 MW; \$/MWh	Block 2 MW; \$/MWh	Block 3 MW; \$/MWh
1	1	10	60	12; $A_1^k \otimes 20$	24; $A_1^k \otimes 50$	24; $A_1^k \otimes 60$
2	2	10	60	12; $A_2^k \otimes 20$	24; $A_2^k \otimes 40$	24; $A_2^k \otimes 70$
3	22	10	60	12; $A_3^k \otimes 20$	24; $A_3^k \otimes 42$	24; $A_3^k \otimes 80$
4	27	10	60	12; $A_4^k \otimes 20$	24; $A_4^k \otimes 44$	24; $A_4^k \otimes 90$
5	23	10	60	12; $A_5^k \otimes 20$	24; $A_5^k \otimes 46$	24; $A_5^k \otimes 75$
6	13	10	60	12; $A_6^k \otimes 20$	24; $A_6^k \otimes 48$	24; $A_6^k \otimes 60$

Bảng 3: Thông tin bảng giá thầu của 3 LSEs

LSE	Nút	Block 1 MW; \$/MWh	Block 2 MW; \$/MWh	Block 3 MW; \$/MWh
1	7	10 ; 100	10 ; 70	10 ; 60
2	15	10 ; 100	10 ; 50	10 ; 20
3	30	10 ; 100	10 ; 60	10 ; 50

Tại mỗi nút của hệ thống, công suất điều tiết tối đa 30MW

Tiến hành thực nghiệm hai trường hợp mô phỏng:

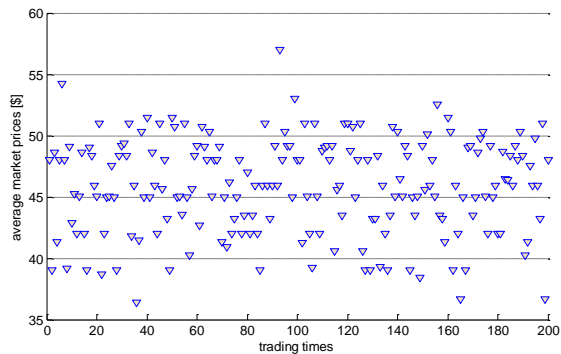
**Trường hợp 01:** Sáu Gencos thay đổi ngẫu nhiên bảng giá chào trong 200 lần giao dịch, không có Gencos nào sử dụng thuật toán ( $A_i^k = 1$ ). Ba LSEs có nhu cầu tải tối thiểu 90MW.

**Trường hợp 02:** Sáu Gencos sử dụng SA-Q learning để tối ưu hóa bảng giá chào trong 200 lần giao dịch, không có Gencos nào sử dụng thuật toán ( $A_i^k$  được lựa chọn trong tập hành động của thuật toán). Ba LSEs có nhu cầu tải tối thiểu 90MW.

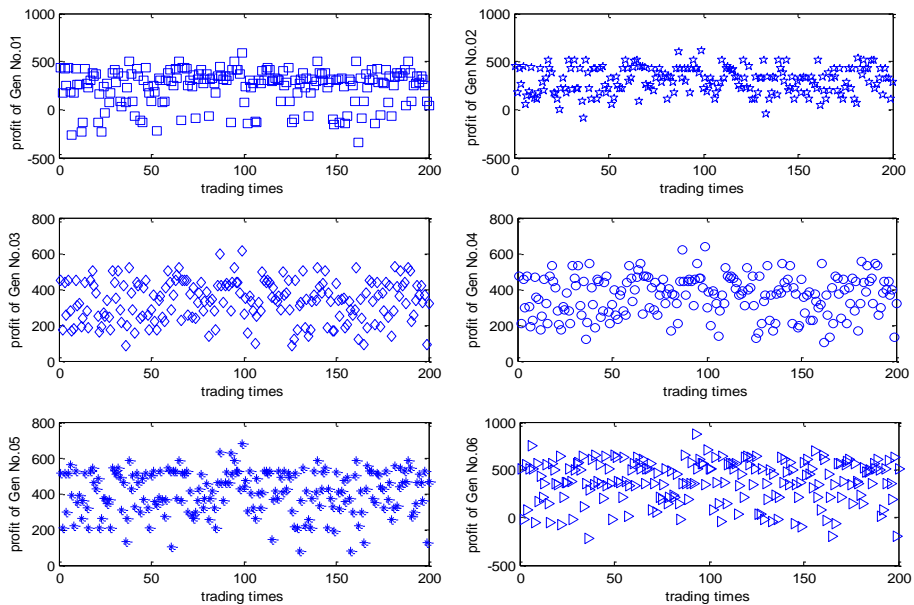
Trong cả hai trường hợp mô phỏng, hai thông số quan trọng được cân nhắc đó là giá trung bình thị trường và lợi nhuận của các GenCos. Kết quả mô phỏng trong trường hợp 1 và trường hợp 2 được trình bày tương ứng ở Hình 8, 9 và Hình 10,11

Kết quả Hình 8 và Hình 9 chỉ ra khi các bảng giá chào và bảng giá thầu được thay đổi dẫn đến giá trung bình trên thị trường thay đổi. Công suất và lợi nhuận của các GenCos luôn biến động. Kết quả chỉ ra trong mô hình thị trường động, giá thị trường biến động liên tục và không ổn định, giá thị trường được dẫn dắt bởi các bảng giá chào và bảng giá thầu. Khuyết điểm lớn của dạng mô hình này cần nhiều cơ chế hơn để kiểm soát.

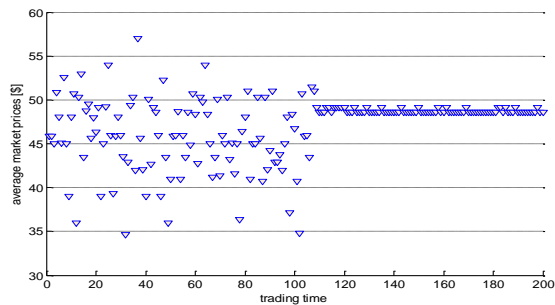
Kết quả tích cực hơn trong trường hợp mô phỏng thứ 2 được trình bày trong hình 10 và hình 11. Khi các GenCos sử dụng thuật toán SA-Q learning thì thông qua 200 lần giao dịch thì giá trung bình trên thị trường và lợi nhuận của các GenCos hội tụ nhanh sau hơn 100 lần giao dịch đầu tiên. Thị trường bắt đầu hoạt động ổn định và yếu điểm của thị trường động được khắc phục.



Hình 8: Trung bình giá điện trong trường hợp 1

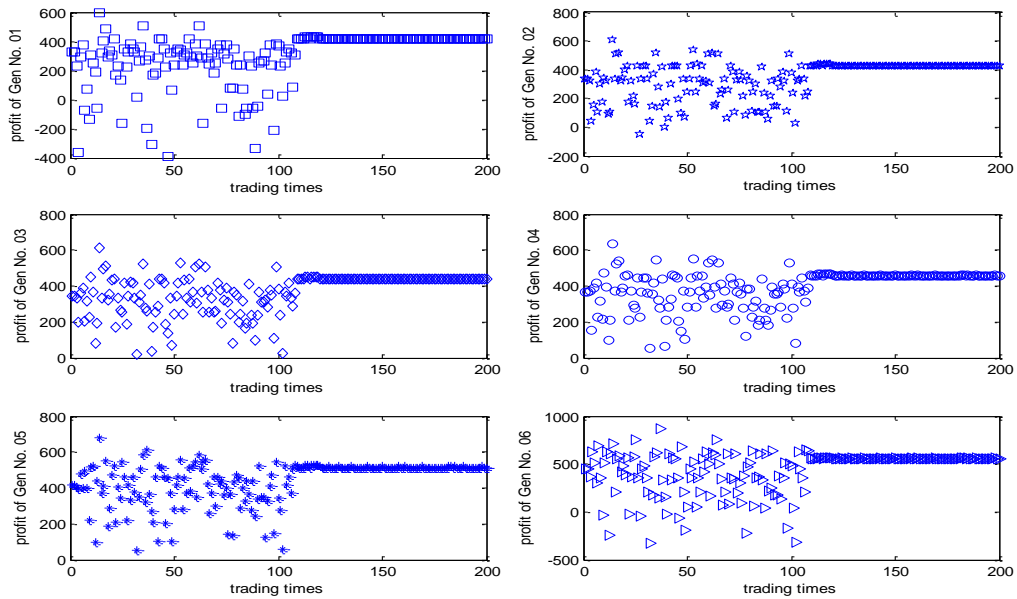


Hình 9: Lợi nhuận đạt được của 06 GenCos trong trường hợp 1



Hình 10: Trung bình giá điện trong trường hợp 2





Hình 11: Lợi nhuận đạt được của 6 GenCos trong trường hợp 2

## 6 KẾT LUẬN

Mô hình thị trường động trên cơ sở cân bằng cung cầu của các đơn vị tham gia là nhu cầu đòi hỏi từ thực tế trong nền kinh tế thị trường. Tuy nhiên, mô hình xuất hiện khuyết điểm lớn đó là sự không ổn định, giá trung bình trên thị trường luôn có biến động. Kết quả mô phỏng cho thấy khi các GenCos sử dụng thuật toán SA-Q learning để tối ưu chiến lược chào giá sẽ làm cho thị trường ổn định và các giá trị hội tụ sau hơn 100 lần biến động giao dịch ban đầu. Mô phỏng thực hiện trên hệ thống tiêu chuẩn IEEE 30 nút cũng được xem là đủ lớn để kiểm chứng các giới hạn và các điều kiện so với hệ thống thực.

Kết quả khích lệ trong nghiên cứu này là cơ sở cho các nghiên cứu sâu hơn trong lĩnh vực thị trường điện thông minh. Một loại thị trường phức hợp và phức tạp hơn khi có sự kết hợp của nhiều loại hình như: Dự trữ năng lượng, năng lượng tái tạo, công suất phản kháng và thị trường hợp đồng song phương. Trong các dạng mô hình thị trường cần phải thiết lập nhiều hơn các ràng buộc như giá carbon thấp, giá chi phí dự phòng và các dịch vụ phụ trợ ... Đây là các chủ đề nghiên cứu cho thị trường điện tương lai./.

## DANH MỤC TÀI LIỆU THAM KHẢO

- [1] Quyết định 26/2006/QĐ-TTg; Quyết định phê duyệt lộ trình, các điều kiện hình thành và phát triển các cấp độ thị trường điện lực tại Việt Nam.
- [2] (2019) Thị trường bán buôn cạnh tranh: “cuộc đua” bắt đầu. [Online]. Available: <https://www.evn.com.vn/d6/news/Thi-truong-ban-buon-dien-can-hanh-2019-Cuoc-dua-bat-dau-6-15-22947.aspx>
- [3] Dong-Joo Kang, Balho H. Kim, Don Hur, Supplier Bidding Strategy Based On Non-Cooperative Game Theory Concepts In Single Auction Power Pools, *Electric Power Systems Research*, vol. 77, 2007, pp. 630 – 636.
- [4] R. W. Ferrero, J. F. Rivera, S. M. Shahidehpour, Application Of Game With Incomplete Information For Pricing Electricity In Deregulated Power Pools, *IEEE Transaction on Power Systems*, vol. 13, n. 1, 1998, pp. 184 – 189.
- [5] Fushuan Wen, A. Kumar David, Oligopoly Electricity Market Production Under Incomplete Information, *IEEE Power Engineering Review*, vol. 21, n. 4, April 2001, pp. 24 – 28.
- [6] Haili Song, Chenching Liu, Jacques Lawarrée, Robert W. Dahlgren, Optimal Electricity Supply Bidding By Markov Decision Process, *IEEE Transactions on Power Systems*, vol. 15, n. 2, May 2000, pp. 618 – 624.
- [7] Charles W. Richter, Jr. Gerald B. Sheblé, Genetic algorithm evolution of utility bidding strategies for the competitive marketplace, *IEEE Transaction on Power Systems*, vol. 13, n. 1, Feb.1998, pp. 256 – 261.

- [8] Charles W. Richter, Jr. Gerald B. Sheblé, Dan Ashlock, Comprehensive Bidding Strategies With Genetic Programming/Finite State Automata, IEEE Transaction on Power Systems, vol. 14, n. 4, Nov.1999, pp.1207 – 1212.
- [9] Y. Y. Hong, S. W. Tsai and M. T. Weng, Bidding Strategy Based On Artificial Intelligence For A Competitive Electric Market, IEE Proceeding Generation Transmission and Distribution, vol. 148, n. 2, pp. 159 – 164 , March 2001.
- [10] Gaofeng Xiong, T. Hashiyama, S. Okuma, An Electricity Supplier Bidding Strategy Through Q-Learning, IEEE Power Engineering Society Summer Meeting, vol. 03, pp. 1516 – 1521, July 2002.
- [11] Chen Haoyong, Yang Yan, Zhang Yao, Realization Of Decision Making Module In Agent-Based Simulation Of Power Markets, Automation of Electric Power Systems on China, vol. 32, n. 20, Oct. 2008, pp. 22 – 26.
- [12] Thanhquy Bach, Jiangan Yao. The SA – Q learning algorithm application to the wholesale power markets. International Journal of Electrical Engineering & Technology, 2012, Vol. 3, No.1 (1-15)
- [13] Maozu Guo, Yang Liu, Jacek Malec. A New Q – Learning Algorithm Based on The Metropolis Criterion. IEEE Transactions on Power Systems. Vol.34, No.5, pp.2140-2143.
- [14] Anastasios G. Bakirtzis, Athina C. Tellidou. Agent-Based Simulation of Power Markets under Uniform and Pay-as-Bid Pricing Rules using Reinforcement Learning, 2006 IEEE Power Systems Conference and Exposition, Atlanta, pp.1168-1173.
- [15] Thanhquy Bach, Jiangan Yao, Shengjie Yang. Fuzzy Q – Learning for Uniform Price Wholesale Power Markets, 2013 International Conference on Communication Systems and Network Technologies, 6-8 April 2013, Gwalior, India.
- [16] Sawan Sen, S. Sengupta, Chakrabart. Alleviation of Line Congestion using Multiobjective Particle Swarm Optimization, International Journal of Electronic and Electrical Engineering, ISSN 0974-2174 Volume 4, Number 1 (2011), pp.123-134.

*Ngày nhận bài: 14/12/2020*

*Ngày chấp nhận đăng: 30/03/2021*