Simulation of the Bidding Behavior of Electricity Retailers Based on the Q-Learning Algorithm

Yuting Xie,1 Lin Tian,1 Jinqing Luo,1 Xiuzhen Hu,1 Jie Qin,1 Kangan Shu,1 and Wenyuan Huang2

1Guangdong Power Exchange Center Co. Ltd., Guangzhou 510000, China
2China Electric Power Research Institute, Nanjing 210003, China

Correspondence should be addressed to Yuting Xie; 604546979@qq.com

Received 14 February 2022; Revised 1 March 2022; Accepted 26 March 2022; Published 28 May 2022

Academic Editor: Zhiguo Qu

At present, the reform of electricity marketization has become a new trend in the development of the world’s electric power industry and electricity retailers have a huge role in further activating power market competition. Simulation of the bidding behavior of electricity sales companies is helpful to analyze the rationality of the transaction behaviors of electricity sales companies and predict the potential risks of market operation. Considering the complex and diverse behaviors of electricity retailers and the immaturity of existing simulation technologies in the power sales market, this paper proposes a Q-learning algorithm-based simulation method for the bidding behavior of electricity retailers. Firstly, we analyze the factors influencing the bidding behavior of electricity retailers, then, we construct a simulation model of electricity retailers’ bidding behavior based on the Q-learning algorithm, and finally, we simulate the bidding process of medium- and long-term trading in provincial power markets by using simultaneous learning from both sides of the generation and consumption. The results show that the behavior of different types of electricity retailers meets the expected level of the model and the simulation method has generalization and extension significance. On this basis, this paper also proposes a bid simulation system architecture based on the bid line simulation model of electricity retailers, which provides a certain reference value for market operation agencies to carry out the corresponding deduction work.

1. Introduction

Since the market-oriented reform of the power industry was launched in the 1970s, the world’s power improvement process has gone through half a century [1, 2]. The power market of each country has gone through three stages: introducing competition into the field of power generation, opening the power market on the power generation side, and selling the power market with competition on both sides. The vigorous development of the electricity market has brought benefits to countries such as the continuous release of market dividends, the overall reduction of electricity costs, and the diversified development of electricity sellers [3–5]. However, after entering the deepwater phase of electricity reform, problems such as dual-track operation of the planned market, inadequate cultivation of the customer-side market, insufficient risk control capability of electricity retailers, and increasing pressure on the operation of electricity retailers have become increasingly prominent. It is of great significance to improve the construction of various countries’ electricity sales market, promote the development of energy Internet and digital economy, and help achieve the low-carbon goals [6–8].

The operation of the electricity market is a long-term dynamic evolutionary process, and the market behavior of electricity sales companies is constantly adjusted, so it is difficult to predict the market operation effect only through theoretical analysis [9]. In addition, electricity sales companies do not have effective means to grasp customers’ electricity consumption curves and most of the retail contracts are signed in the medium- and long-term market, so the behavior of electricity sales companies in medium- and long-term transactions needs to be simulated by means of simulation [10, 11].
In order to simulate the medium- and long-term trading behavior of electricity sales companies, it is necessary to analyze the operating conditions of electricity sales companies and clarify their profitability models. Huang [12] simulates the oligopolistic competitive market and the equilibrium market by constructing a coalition pricing game model and gives suggestions to reasonably set the electricity sales limit and strengthen the market behavior cultivation of electricity retailers; Song et al. [13] summarized and analyzed the current situation of research on electricity retailers’ power purchase and sales decisions from three aspects: prediction of uncertainty parameters in the power sales market, power purchase decisions of electricity retailers, and pricing strategies of electricity retailers, and discusses the future development. Some scholars [14–16] considered the customer’s independent choice behavior and designed the pricing strategy of retail electricity packages for electricity sales companies to enhance the ability of electricity sales companies to attract new customers and maintain the stickiness of old customers; Wang et al. [17] introduced the concept of value-added electricity services, established a sales strategy model of electricity products and value-added services considering customer network externalities and customer preferences, and conducted a dual-oligopoly electricity sales market. The simulation analysis was conducted.

Secondly, a scientific and reasonable simulation learning model of electricity sales company behavior needs to be constructed to reasonably quantify the simulation effects and provide decision support for electricity sales companies’ market transactions. Some scholars [18–20] use the VRE-learning reinforcement learning algorithm, multiagent double-DQN, algorithm and improved reinforcement learning automatons to simulate and learn the generators’ bidding behavior, which provide a reference for the study of electricity sales company behavior simulation. Wang et al. [21] combined the price quota curve and market clearing tariff to build a model of electricity sales company bidding behavior and simulated. A bidding model of electricity retailers under deviation assessment mechanism is established based on the robust optimization algorithm, and the optimal strategy is solved by the genetic algorithm, and the simulation shows that the model can combine economy and risk resistance [22]. Dou et al. [23] establishes an optimization model of power purchase and sale for generation-qualified electricity retailers and independent electricity retailers based on the segmented offer rule and revealed the profit changes of different qualified electricity retailers in the market game through simulation.

The abovementioned analysis shows that there are two shortcomings in the existing literature: firstly, the research on the bidding strategy of electricity retailers mainly focuses on the profit maximization of electricity retailers and there is a lack of modeling of the behavior of electricity retailers with different asset backgrounds and different business objectives; secondly, the simulation of the medium- and long-term power sales market mainly focuses on the pricing strategy of the electricity retailers themselves and there is no dual-side bidding from the provincial market perspective. The design of the simulation is not designed from the provincial market perspective, and the simulation results have some limitations.

In view of this, this paper investigates the simulation model of medium- and long-term trading behavior of electricity retailers. Firstly, we analyze the factors influencing the behavior of electricity retailers, based on which we propose a simulation model of electricity retailers’ bidding behavior based on Nash equilibrium theory and Q-learning algorithm, and then, we construct a provincial medium- and long-term market simulation scenario to simulate the market game situation in order to optimize the bidding strategy of electricity retailers themselves. Finally, the validity of the bidding model is verified by simulating a bilateral bidding case in the provincial medium- and long-term market. This paper also proposes a bid simulation system architecture based on the bidding line simulation model of the electricity retailer in order to reduce the decision risk of the electricity retailer.

2. Analysis on Influencing Factors of Bidding Behavior of Electricity Retailers

After the retail market was liberalized, power generation enterprises, incremental distribution network enterprises, independent investors, energy companies, equipment manufacturers, and other enterprises from different industries and backgrounds have registered as electricity retailers, with different business objectives. With the goal of “carbon neutral,” the transformation of the new power system with new energy as the mainstay has accelerated and the demand-side participation in the market, such as interruptible load, customer-side energy storage, and distributed photovoltaic, has become increasingly demanding [24, 25]. At the same time, most of the demands for electricity consumption and purchase from a large number of customers under the new electricity reform environment will be mapped to the electricity retailers as agents. As a result, the behavior of electricity retailers in the wholesale market is bound to show complexity and variability.

In this paper, the factors influencing the bidding behavior of electricity sellers are divided into three dimensions: market rules, market environment, and sellers’ own factors.

2.1. Market Rules

2.1.1. Quotation Rules

(i) Single-segment quotation: each trading session only allows market members to declare one electricity quantity and tariff, and if they cannot win the bid, they will fail to bid directly, so electricity retailers will be extremely cautious in quoting.

(ii) Multisegment offer: each trading session allows market members to declare multiple power and tariffs, commonly three segments, six segments, seven segments, ten segments, etc. Under the multisegment offer rule, most electricity retailers will bid high in the first few segments to ensure that they win, and then, they will consider quoting low prices for some of the power to make more profit.
2.1.2. Deviation Assessment Rules

(i) Assessment object: for the deviation assessment object of the electricity sales market, the current rules vary around the world, including three kinds of overall assessment of the electricity sales company, assessment of users, and retail contracts in their own agreement.

(ii) Assessment methods: the severity of the assessment methods affects the aggressiveness of the bidding behavior of electricity retailers to a certain extent, the most flexible is the rolling adjustment method, and most markets adopt the assessment fee within a certain deviation range, and companies with large deviations face the penalty of market access disqualification or even contract termination within a certain market.

2.2. Market Environment

2.2.1. Supply and Demand Ratio

(i) Supply is greater than demand: competition is fierce on the generation side, and the electricity retailer adopts a low-price strategy.

(ii) Supply is less than the demand: competition is fierce on the demand side, and the electricity retailer quotes high prices to ensure winning bids.

(iii) Supply and demand are equal: competition is fierce on the supply and demand side, and the electricity retailer selects its offer according to its own characteristics.

2.2.2. Supply Side

(i) Coal price: as the source of the electricity production chain, the price of coal has a direct impact on the price of electricity and the signal can often only be transmitted to the intermediate power sales link but not to the user side in a timely manner. In 2021, for example, coal prices rose sharply, making domestic thermal power companies difficult to operate and less willing to generate electricity, resulting in large losses for electricity retailers in Guangdong and Shanxi [26].

(ii) New energy output: new energy output is severely lacking due to extreme weather, resulting in a major blackout in Texas, USA, on February 15–19, 2021, with spot market prices exceeding $9,000/(MW-h). In the future, in the new power system with new energy as the main source, the impact of new energy on the power market price will bring some risk to the operation of electricity retailers [27, 28].

2.3. Electricity Retailers’ Own Factors

2.3.1. Asset Background

(i) Power plant background type: the main goal is to sell power from the plant.

(ii) User background type: the goal is saving the cost of buying electricity.

(iii) Pure intermediary type: the goals are survival and profitability.

(iv) Energy service background type: the goal is to supplement the comprehensive energy service capability with electricity.

(v) Owning distribution network asset background type: through coordination of resources, the goal is to complete the task of selling electricity.

2.3.2. Operational Objectives

(i) Expansion type: to attract customers and expand the market scale.

(ii) Profit oriented: aiming at pure profitability.

(iii) Robust: less fluctuation in bidding behavior.

2.3.3. Market Share

(i) Electricity retailers with a large market share can play the role of price setters, while those with a very small market share can only act as price takers.

2.3.4. Other Market Transactions. The annual market, monthly market, and spot market influence each other.

2.3.5. The Situation of End Retail Contracts

(i) Fixed spread-type contract: the electricity seller and the customer agree on a fixed spread, and the customer enjoys preferential electricity rates based on this spread, which is not affected by market price fluctuations.

(ii) Guaranteed plus share-type contract: a guaranteed tariff-and-share ratio is agreed between the electricity seller and the customer; when the liquidation price of the monthly centralized trading is higher than the guaranteed price, both parties trade at the contract tariff, and when the liquidation price is lower than the guaranteed price, both parties share the profit for the excess profit [29].

3. Simulation Model of the Bidding Behavior of the Electricity Retailer

In order to reflect the bidding behavior of each electricity retailer in the electricity market more realistically, this chapter constructs a simulation scenario of electricity retailers' bidding behavior based on comprehensive consideration of market rules, market environment, and electricity retailers' own factors and analyzes the typical electricity retailers' bidding behavior strategies and benefits by combining the Nash equilibrium theory. Based on this, the Q-learning algorithm is used to train the bidding strategies of the current market members.
3.1. Simulation Scenario Analysis of Bidding Behavior of Electricity Retailers Based on the Nash Equilibrium Theory.
In the electricity market, any one market member does not know the cost of the other members, so the trading model of the market members is actually a noncooperative game with incomplete information, i.e.,

\[ G = \{ M, A, P, U \}, \]
\[ M = \{ m_1, \ldots, m_n \}, \]
\[ A = \{ A_1, \ldots, A_i, \ldots, A_k \}, \]
\[ A_i = \{ a_{i1}, \ldots, a_{ij}, \ldots, a_{ik} \}, \]
\[ P_i = \{ p_{i1}, \ldots, p_{ij}, \ldots, p_{ik} \}, \]
\[ \text{s.t.} \quad 0 \leq p_{ij} \leq 1, \]
\[ \sum_{j=1}^{k} p_{ij} = 1, \tag{3} \]

where \( M \) is the set of market members, \( A \) is the set of bidding strategies, \( P \) is the set of strategy selection probabilities, \( U \) is the set of market members’ operational target returns, and \( k \) is the total number of market members’ bidding strategies.

Suppose that the number of market members is \( n \). In the game process, market member \( i \) selects the corresponding strategy \( j \) according to the probability of strategy distribution \( p_{ij} \), and after the market is cleared, the revenue of the strategy is calculated according to the transaction results. However, the market is an organic whole and the interests of each market member are in conflict; in order to better solve this problem, this paper introduces the Nash equilibrium for strategy optimization.

The Nash equilibrium (Nash equilibrium) describes the scenario simply summarized as follows: in a game process, when a member chooses a certain determined strategy, the rest of the members cannot benefit from the change in strategy. That is, after repeated games of market members, the market as a whole will reach Nash equilibrium, at which time the member \( i \) will choose a certain definite offer strategy \( a_i \) (\( a_i \in A_i \)), regardless of how the rival firm offers, at which time there are

\[ u(a_i) \geq u(a_j), \quad (i \in [1, k]). \tag{4} \]

3.2. Bidding Behavior Strategy and Benefit Analysis of Typical Electricity Retailers. This section classifies existing electricity retailers into four types based on their corporate size and development strategies, namely, expansionary electricity retailers with generation background, service sales companies with integrated energy background, robust electricity retailers with large corporate background, and profit sales companies with small-scale private background. The game strategy and revenue objective model of typical electricity retailers are proposed considering the retail contract data of electricity retailers, the penalty power purchase factor of unsuccessful electricity retailers, and other market winning factors and combined with the Nash equilibrium theory in the previous section.

3.2.1. Power Generation Expansion Type of the Electricity Retailer. This type of electricity retailer has upstream resources, obvious competitive advantage, and low profit pressure. The main purpose of the initial stage of market construction is to expand and attract customers by playing low prices, so its bidding behavior is to sign more sales contracts with customers by flexible and low prices and then obtain power to perform with higher offers in centralized bidding and for the deviation part of power purchase needs to be purchased in the market by higher prices. The target model is as follows.

\[ \max \quad u = (\bar{P}_{con} - \bar{c}_g) \times Q_{clr} - (Q_{bid} - Q_{clr}) \times (\bar{P}_{pu} - \bar{P}_{con}), \tag{5} \]

where \( U \) is the revenue of the electricity retailer, \( \bar{P}_{con} \) is the average retail price signed by the electricity retailer and users, \( \bar{c}_g \) is the average power generation cost of power generation business, \( P_{con} \) is the market clearing price, \( Q_{clr} \) is the clearing electricity of the electricity retailer, \( Q_{bid} \) is the bidding electricity of the electricity retailer, and \( P_{pu} \) is the deviation electricity price of the electricity retailer.

3.2.2. Comprehensive Energy Background Service Electricity Retailer. With the rapid development of the energy Internet and the continuous promotion of the distribution-side reform, this type of electricity retailer is also rapidly emerging as a comprehensive energy service provider in industrial parks; the sale of electricity is only part of its business, together with the microgrid, distributed generation, and other resources held; its bidding behavior is manifested in the form of higher price difference to ensure that it attracts park users; the game of low price motivation is not strong enough, not in accordance with the contractual agreement to purchase the electricity and the need to use the purchase deviation price to purchase electricity. The target model is as follows.

\[ \max \quad u = (P_{con} - P_{clr}) \times Q_{clr} - (Q_{bid} - Q_{clr}) \times (P_{pu} - P_{con}). \tag{6} \]

3.2.3. Large Enterprise Background Stable Electricity Retailer. As a large enterprise specially established electricity retailer, the main service object of this type of electricity retailer is the enterprise itself, so its bidding game behavior is to ensure a certain profit premise through the market transactions to reduce the enterprise’s grid agency power purchase cost.

\[ \max \quad u = (P_{cat} - P_{clr}) \times Q_{clr} \times \lambda(0 < \lambda < 1) \]
\[ \text{s.t.} \quad Q_{clr} \times (P_{con} - P_{clr}) \times (1 - \lambda) \geq P_{exp}, \tag{7} \]

where \( P_{cat} \) is the power grid agent purchase price of the enterprise, \( \lambda \) is the proportion of the electricity consumed.
by the enterprise in the transaction amount, $P_{\text{con}}$ is the contract price signed with other users, and $P_{\text{exp}}$ is the lowest profit condition.

3.2.4. Small-Scale Private Background Profit Electricity Retailer. This type of electricity retailer does not have industry experience and does not have the advantage of power system resources and information, so it can only make profit by earning the difference between purchase and sale, and is willing to play the game of low price more strongly. Therefore, their market behavior is to sign a guaranteed plus share contract with customers to protect their own profits while reducing the market risk that they bear and attracting customers.

$$\begin{align*}
\text{If } (P_{\text{clr}} \leq P_{\text{gua}}), \\
\max \ u = (P_{\text{con}} - P_{\text{clr}}) \times \theta \times Q_{\text{clr}} - (Q_{\text{bid}} - Q_{\text{clr}}) \\
\times [ (P_{\text{pu}} - P_{\text{con}}) + (P_{\text{con}} - P_{\text{clr}}) \times (1 - \theta) ]. \\
\end{align*}$$

(8)

$$\begin{align*}
\text{If } (P_{\text{clr}} > P_{\text{gua}}), \\
\max \ u = (P_{\text{con}} - P_{\text{clr}}) \times Q_{\text{clr}} - (Q_{\text{bid}} - Q_{\text{clr}}) \times P_{\text{pu}},
\end{align*}$$

(9)

where $P_{\text{gua}}$ is the guaranteed price and $\theta$ is the share ratio of the electricity retailer.

3.3. Simulation Model of Bidding Behavior of Electricity Retailers Based on the Q-Learning Algorithm. In order to better simulate the noncooperative game phenomenon in

Table 1: The behavior of market members simulates the Q-score table.

| State $s$ | Strategy $a_1$ | $a_2$ | ...... | $a_k$ |
|-----------|----------------|-------|-------|-------|
| $s_1$     | $Q(s_1, a_1)$  | $Q(s_2, a_1)$ | ...... | $Q(s_k, a_1)$ |
| $s_2$     | $Q(s_1, a_2)$  | $Q(s_2, a_2)$ | ...... | $Q(s_k, a_2)$ |
| ......     | ......           | ...... | ...... | ...... |
| $S_n$     | $Q(s_n, a_1)$  | $Q(s_n, a_2)$ | ...... | $Q(s_n, a_k)$ |

Figure 1: Flowchart of behavior simulation of electricity retailers based on the Q-learning algorithm.
the real market, market members are required to continuously learn to update their bidding strategies according to the target revenue and market state changes and the whole process can be described as a cyclic process in which market members update their choice probabilities and thus change their bidding strategies to gain revenue and interact with the market, which is a typical Markov decision process [30–32], so this paper adopts the Q-learning algorithm to realize the self-learning of market members’ bidding strategies.

The Q-learning algorithm is a value-based algorithm in reinforcement learning, where Q is the action utility function, and a Q-table is eventually obtained through continuous learning, and the algorithm process is described in the literature [33]; on the other hand, market members explore back and forth between the offer strategy with high returns and the residual strategy to prevent premature entry into the local optimum and ensure the optimality of the final decision. The learning process is shown in Figure 1.

The Q-learning-based simulation process for the behavior of electricity retailers includes six steps.

(i) Set the simulation parameters, including the number of generators, the number of electricity retailers, the ratio of supply and demand, the type of electricity retailers, the bidding space, the number of bidding strategies, and the exploration probability p

(ii) Generate and initialize the Q value table from the market state s and bidding strategy a, as shown in Table 1; determine the maximum number of learning rounds N

(iii) Randomly generate the selection probability r. When r is greater than p, market members will choose any offer strategy and expand the Q value table; when r is less than p, market members will choose the best offer strategy based on the existing Q value table

(iv) All market members’ offer data will be integrated and recorded, and then, the clearing calculation service will be invoked to get the market clearing tariff and clearing power

(v) Calculate the revenue \( u_i(t) \) of the electricity retailer according to the revenue model proposed in Section 4.2, and calculate the Q value of the corresponding state strategy according to equations (10) to (11), and then, update the Q value table and the strategy selection probability table

\[
Q_{t+1}(s, a) = \begin{cases} 
(1 - \alpha)Q_t(s, a) + \alpha(u'_i + \gamma H), & s = s_t, a = a_{ij}, \\
Q_t(s, a), & \text{otherwise},
\end{cases}
\tag{10}
\]

\[
H = \max Q_t(\tilde{s}, a),
\tag{11}
\]

\[
P_i(s, a) = \frac{\exp\left[Q_t(s_t, a_{ij})\right]}{\sum_{j=1}^{K}\exp[Q_t(s, a)]},
\tag{12}
\]

where \( \alpha \) is the learning rate, \( \gamma \) is the discount factor, \( u'_i \) is the T-round behavior gains, and \( H \) is the historical maximum Q value of the new state \( \tilde{s} \). It can be seen in the formula that the greater the learning rate, the smaller the discount factor and the more market members pay attention to the current income; the smaller the learning rate, the larger the discount factor and the more market members rely on historical experience

(vi) If the quotation round is equal to \( T \), determine the optimal quotation strategy set and end the learning; otherwise, repeat steps (2) to (4)

### 4. Example Verification

In order to verify the validity of the model built in the previous section, this chapter takes a medium- and long-term monthly direct trading in an electricity market as an
example and builds a provincial electricity market simulation and projection environment with the simulation software MATLAB 2019b.

4.1. Parameter Setting. In this section, the simulation results of two different scenarios will be used to analyze the behavior of different types of electricity retailers under different offer rules. The market environment parameters are shown in Table 2, where market members G1–G100 are cost-offer-type generators, market members S1–S50 are generation background expansion-type electricity retailers, electricity retailers S51–S100 are integrated energy background service-type electricity retailers, electricity retailers S101–S150 are large enterprise background robust electricity retailers, and electricity retailers S151–S200 are small private background profit-oriented electricity retailers. The market membership parameters are shown in Table 3.

The Q-learning algorithm sets the learning rate $\alpha$ to 0.05, the discount factor $\gamma$ to 0.95, and the initial Q value to 0.

4.2. Result Analysis

4.2.1. Comparative Analysis of Quotation Curves before and after Learning by Electricity Retailers. The offer curves before and after learning of the two scenarios are shown in Figures 2 (a)–2(d). As can be seen in Figure 2, after learning, all the declared power by the electricity retailers in the two scenarios are sold, the highest offer of the generators is concentrated around the market clearing price, and all the bids on both sides of the offer are won, which is in line with the expected effect after the market reaches Nash equilibrium under the market environment where supply exceeds demand, indicating that the learning algorithm is real and effective. Table 4 shows the learning data of the Q value and strategy selection probability of market members G1 and S1 in the 1st, 100th, 300th, and 1000th rounds under the three-stage offer scenario.

4.2.2. Comparative Analysis of Clearing Prices under Two Scenarios. The trend of the clearing price during the learning
process of the two scenarios is shown in Figures 3(a) and 3(b). As can be seen in Figure 3, the market clearing price under the single-segment offer method is stable at 370 (¥/MWh) from 248 rounds, while the market clearing price under the three-segment offer method is stable at 342.5 (¥/MWh) until 889 rounds, with a large price difference between the two.

This shows that (1) the single-segment offer method has a high market risk and not winning the bid will lead to a "total loss" for market members, which constrains them to adopt a more rational and cautious offer strategy and (2) the three-segment offer method allows market members to use the last segment of the offer to play a trial game and the market is highly competitive and the external conditions of the demand-side market. Therefore, it is necessary to set a lower limit for each segment of the offer in the market rules.

4.2.3. Analysis of Bidding Behaviors of Electricity Retailers. At present, the quotation rules of domestic provincial electricity markets are dominated by multisegment quotation methods. In order to explore the diversity of quotation strategies of different types of electricity retailers, this paper conducts a detailed analysis of the third segment declaration behaviors of electricity retailers under three segment quotation method scenarios. Figures 4(a)–4(d) show the third-segment offer strategy learning diagrams of electricity retailers S1, S100, S150, and S200, respectively. As can be seen in Figure 4, (1) the offer strategy of S1, a power generation background
expansion type of electricity retailers, is locked at 355 ¥/MWh earlier, which shows that the primary purpose of this type of electricity retailer is to obtain more power to support its expansion behavior under the premise of certain profitability; (2) the offer strategy of S100, a comprehensive energy background service type of the electricity retailer, is concentrated at 350 ¥/MWh to ensure the winning power at a relatively high offer, which is in line with its expected behavior of attracting park customers; (3) the bidding strategy of S150 is steadily lowered as the market clearing price decreases, and its behavior is steady, but it is sensitive to the market while satisfying its own demand for electricity; and (4) the bidding behavior of S200 is relatively aggressive, and its bidding strategy follows the changes in the clearing price, and it keeps trying to get lower prices to obtain more electricity and profits at lower prices.

5. A Bidding Simulation System Architecture Based on the Bidding Bank Simulation Model of Electricity Retailers

By applying a medium- and long-term transaction data of electricity market users to validate the bidding behavior simulation model of electricity retailers, the results show that the model can truly reflect the actual bidding behavior of electricity retailers; therefore, this section designs the bidding behavior model system of electricity retailers based on this simulation model to provide reference for electricity retailers to participate in the electricity market bidding. Figure 5 shows the schematic diagram of the business architecture of the electricity retailer bidding behavior simulation software.

The electricity retailer bidding behavior simulation software can be divided into five parts according to the business process: simulation scenario setting → simulation offer parameter setting → offer simulation → clearing simulation → simulation analysis. Among them,

(i) The simulation scenario setting is the basic setting, including market member model management, electricity retailer behavior analysis, and case creation functions

(ii) The simulation offer parameter setting is used to set input parameters for medium- and long-term transaction simulation cases, such as offer range, electricity retailer’s bidding strategy library, bidding algorithm parameters, and electricity retailer’s operation target. This process is designed according to the parameter setting content in chapter 3.
The offer simulation application, also known as the proxy bidding application, is based on the basic setup parameters and completes the calculation of the bidding proxy by invoking the responsive bidding algorithm. The bidding algorithm supports strategy self-learning and offer self-selection. This process is designed based on the content of the behavioral simulation model of the electricity retailer in chapter 2.

Outbound simulation application mainly realizes the outbound rule setting, outbound algorithm invocation and calculation, and outbound electricity retailer revenue calculation. The revenue of the electricity retailers can be calculated based on the typical revenue model of the electricity retailer proposed in Section 2.2, and the market clearing rules and clearing method are set according to the actual situation of the electricity retailer.

Simulation analysis applications, supporting the analysis of simulation results from multiple dimensions, are the trend of change of electricity retailer’s revenue, the evolution of electricity retailer’s offer strategy, and the change of market clearing price.

6. Conclusion

Aiming at the diversity of behaviors of electricity sales companies in medium- and long-term centralized bidding transactions, this paper proposes a simulation method of bidding behavior of electricity sales companies based on the Q-learning algorithm. The behavior difference modeling and simulation of the power generation background expansion-type power sales company, the comprehensive energy background service power sales company, the large enterprise background stable power sales company, and the small-scale private background profit-based power sales company are carried out. The final convergent quotation is consistent with the bidding strategy of the electricity sales company and reaches the Nash equilibrium, which verifies the effectiveness of the method. On this basis, a bidding behavior bidding simulation system architecture of electricity sales companies is designed, which covers the simulation scene setting, simulation quotation parameter setting, quotation simulation, clearing simulation, and simulation analysis functions. This architecture has a certain promotion value.

While some research results are achieved, there are also shortcomings in this paper, such as only four types of factors influencing the bidding behavior of electricity retailers, and their bidding behavior strategies are analyzed but the factors influencing the bidding offers of electricity retailers in the actual power market and their performance behaviors still need to be further explored. At the same time, based on the design idea of this paper, the bidding behavior of various types of electricity retailers can be simulated more accurately and comprehensively by extending the operation objective function of electricity retailers and further subdividing the behavior of electricity retailers.
Data Availability
No data were used to support this study.

Conflicts of Interest
There is no conflict of interest regarding the publication of this paper.

Acknowledgments
This paper is supported by the project “science and technology project of Guangdong Power Grid Co. Ltd.” (GDKJXM20185364).

References
[1] S. Guo, W. Zhang, and X. Gao, “Business risk evaluation of electricity retail company in China using a hybrid MCDM method,” Sustainability, vol. 12, no. 5, p. 2040, 2020.
[2] M. Zeng, Y. Yang, L. Wang, and J. Sun, “The power industry reform in China 2015: Policies, evaluations and solutions,” Renewable and Sustainable Energy Reviews, vol. 57, pp. 94–110, 2016.
[3] P. You, S. Guo, H. Zhao, and H. Zhao, “Operation performance evaluation of power grid enterprise using a hybrid BWM-TOPSIS method,” Sustainability, vol. 9, no. 12, p. 2329, 2017.
[4] Y. He, J. Jiao, R. Chen, and H. Shu, “The optimization of Chinese power grid investment based on transmission and distribution tariff policy: a system dynamics approach,” Energy Policy, vol. 113, pp. 112–122, 2018.
[5] X. Peng and X. Tao, “Cooperative game of electricity retailers in China’s spot electricity market,” Energy, vol. 145, pp. 152–170, 2018.
[6] Q. Peng, X. J. Zhou, R. Yang, and W. T. Jia, “Bidding strategy of comprehensive energy based power selling company participating in electricity market competition under ubiquitous environment of internet of things,” Power System Technology, vol. 43, no. 12, pp. 4337–4343, 2019.
[7] State Grid News, “Beijing Electric Power Trading Center Released the Trading Information of Electricity Market in 2020,” 2021.
[8] Guangzhou Electric Power Trading Center, “2020 South Regional Trans-regional and Inter-provincial Electricity Market Operation Report,” 2021.
[9] J. R. Andrade, J. Filipe, M. Reis, and R. J. Bessa, “Probabilistic price forecasting for day-ahead and intraday markets: Beyond the statistical model,” Sustainability, vol. 9, no. 11, p. 1990, 2017.
[10] J. Zhang, Y. Zheng, M. Yao, H. Wang, and Z. Hu, “An agent-based two-stage trading model for direct electricity procurement of large consumers,” Sustainability, vol. 11, no. 18, p. 5031, 2019.
[11] S.-V. Oprea, A. Băra, A. Uţă, A. Pirjan, and G. Caruţaşu, “Analyses of distributed generation and storage effect on the electricity consumption curve in the smart grid context,” Sustainability, vol. 10, no. 7, p. 2264, 2018.
[12] Y. X. Huang, “Research of the oligarchic game of monopoly and competition on the electricity market after introducing competition into the retail side,” Journal of Southwest University (Natural Science Edition), vol. 43, no. 5, pp. 142–151, 2021.
[13] Y. H. Song, G. Wang, H. Cai, C. W. Gao, and T. Chen, “Research review of electricity purchase and sale trading strategies,” Power Demand Side Management, vol. 22, no. 6, pp. 85–89, 2020.
[14] R. Li, Z. Wang, C. Gu, F. Li, and H. Wu, “A novel time-of-use tariff design based on Gaussian Mixture Model,” Applied Energy, vol. 162, pp. 1530–1536, 2016.
[15] J. Hou, Z. Lin, L. Yang, Y. Ding, K. Luan, and B. Yang, “Design of electricity plans for industrial and commercial customers oriented to active demand response on power demand side,” Automation of Electric Power Systems, vol. 42, no. 24, pp. 11–21, 2018.
[16] E. Lu, P. Bie, H. H. Wang, Q. Chen, W. X. Li, and W. Z. Gu, “Pricing strategy design of electricity retail pricing package considering self-selectivity of users,” Automation of Electric Power Systems, vol. 44, no. 19, pp. 177–184, 2020.
[17] Z. Y. Wang, W. H. Zhao, P. Gao, W. Z. Zhao, and K. L. Zhong, “Sale strategy model of electricity value-added services considering network externalities,” Power System Technology, vol. 42, no. 7, pp. 2213–2222, 2018.
[18] H. Feng, Z. L. Yang, Y. X. Zheng, F. Ye, X. Zhang, and X. Shi, “Intelligent agent based bidding simulation methods of multi-input decision factors for power suppliers based on intelligent agent,” Automation of Electric Power Systems, vol. 42, no. -appeec, pp. 72–77, 2018.
[19] Y. Gao, Y. Li, R. R. Cao, N. F. Li, and M. Z. Gao, “Simulation of generators’ bidding behavior based on multi-agent double DQN,” Power System Technology, vol. 44, no. 11, pp. 4175–4183, 2020.
[20] G. Q. Jia, S. J. Chen, Y. Y. Li, Z. Yan, and C. K. Xu, “Learning automata based bidding strategy for power suppliers in incomplete information environment,” Automation of Electric Power Systems, vol. 45, no. 6, pp. 133–139, 2021.
[21] D. H. Wang, Y. Wang, R. Z. Cao, and K. F. Zhang, Price Quota Curve Based Optimization of Bidding Strategy for Electricity Retailers, Power System Automation, 2021.
[22] C. H. Sun, P. Li, R. F. Shi, and T. B. Zhang, “Robust bidding strategy for power retailer considering deviation assessment in electricity market,” Distribution & Utilization, vol. 38, no. 7, pp. 72–78, 2021.
[23] X. Dou, P. Zhang, J. A. Li, and J. Wang, “Analysis of power purchase-sale strategy of power retail companies with different establishment qualifications,” Proceedings of CSEE, vol. 40, no. S1, pp. 181–187, 2020.
[24] L. Yin, J. Liu, H. Gao, S. Ning, and Y. Li, “Study on bidding strategy of comprehensive power retailer under multiple user-price mechanisms,” Power System Technology, vol. 42, no. 1, pp. 88–97, 2018.
[25] Y. S. He, J. F. Yang, and F. X. Chen, “Research on operation model of electricity retail company based bi-level optimization,” Journal of Electric Power Science and Technology, vol. 34, no. 3, pp. 71–77, 2019.
[26] D. Q. Li, J. R. Gong, Z. L. Li, Y. Gao, and F. S. Wen, “Investigation of the thermal performance of a novel thermosyphon combined with fluidized bed heat transfer technology,” Journal of Electric Power Science and Technology, vol. 374, no. 1, pp. 40–48, 2020.
[27] C. Hu, S. H. Du, J. Su, G. Y. Tong, and M. Z. Wang, “Preliminary research of trading approach and management modes of...
Chinese electricity retail companies under new electricity market reform,” *Power System Technology*, vol. 40, no. 11, pp. 3293–3299, 2016.

[28] Y. Zhang, G. L. Xie, Q. Zhang, X. Y. Han, R. Zhu, and J. Zhang, “Analysis of 2.15 power outage in Texas and its implications for the power sector of China,” *Electric Power*, vol. 54, no. 4, pp. 192–198+206, 2021.

[29] L. Tang and J. C. Liu, “Multi-target electricity procurement and sale strategies modeling of electricity retailers under different retail contract modes,” *Distribution & Utilization*, vol. 36, no. 2, pp. 48–55, 2019.

[30] H. H. Yu, S. H. Lin, J. Q. Zhu, and H. W. Chen, *On-Line Optimization of Micro Grid Based on Deep Reinforcement Learning*, Electrical Measurement and Instrumentation, 2021.

[31] G. D. Guo and Y. F. Gong, “Real-time automatic control algorithm of micro-grid energy management system based on deep reinforcement learning in electricity market environment,” *Electrical Measurement and Instrumentation*, vol. 58, no. 9, pp. 78–88, 2021.

[32] P. Lv, S. Y. Pan, J. Xu, and T. S. Li, “A network traffic classification algorithm based on machine learning,” *Journal of Guangxi University (Natural Science Edition)*, vol. 44, no. 6, pp. 1650–1657, 2019.

[33] T. Q. Liu, D. Han, Y. D. Wang, and X. T. Dong, *Strategies of Market Game Behavior of Virtual Power Plants Based on Q-Learning with Augmented Lagrange Function*, Power System Technology, 2021.