Identification of market power abuse in China’s electricity market

Ziyun Cheng¹, Hao Li¹, Huihui Xu¹, Li Dong²*, Hongyuan Tao³, Huichun Hua², Shenghua Wang²

¹Economic and technological research institute of State grid power company of Gansu province, Lanzhou, Gansu province, China
²State key laboratory of the new energy power system (North China Electricity Power University), Baoding, Hebei province, China

*Corresponding author’s e-mail: DongLi@ncepu.edu.cn

Abstract: The goal of this paper is to find a method to identify violations, which is suitable for China’s electricity market. Therefore the key problems in the process were discussed, that is, the trading rules and marketization degree which were different from those of foreign countries. So this article has done the following works to solve the problems. Firstly, this paper classifies and compares the identification indexes of market power abuse, and analyzes the advantages and disadvantages of each index, based on which an index system is proposed. Secondly, An intelligent identification algorithm based on support vector machine (SVM) is introduced and improved to solve the data imbalance problem. Also, the index system and the improved SVM were combined to cover the economic principle as well as improve the identification efficiency at the same time. Taking the simulation data as an example, the rate of violation recognition reaches 100%, which indicates that the method proposed in this paper is feasible and effective.

1. Introduction

Since the release of the Opinions on Further Deepening the Power System Reform on March 15, 2015, China has begun to comprehensively promote a new round of power system reform. As all aspects of reform are flourishing, violations such as abuse of market power emerge due to the lack of regulatory policies. However, the current violation identification in China mainly rely on inefficient post-bidding expert judgments. So with the further increase in transaction volume, the electric market urgently needs an efficient method for identifying violations. Among all the violations, the market power abuse is the most important factor affecting the transaction price so that it needs to be addressed first.

From the perspective of economics, market power refers to the ability of the provider of a product or service to continuously influence or manipulate the market price, making it always higher than the market price of the perfect competition market. It is a very complicated problem to judge whether a generator has abused market power or not. The existing literature can be mainly divided into three categories: review, principle, and identification methods, among which the identification methods can be divided into index method, simulation method, and intelligent identification. Literature [1-4] are review articles, which summarize the methods for monitoring violations in the power market and the experience of foreign market power monitoring. The principle of market power is discussed in Literature [5-8], of which the literature [5] describes the principles and conditions for the formation of market power. About the identify -cation method, literature [9-13] use the index method, among which
literature [9-12] respectively introduce the market power monitoring indicators of India, the Baltic sea, Singapore and the UK, literature [13] proposes three new indexes to evaluate the market power of new energy. The second identification—simulation method is introduced in literature [14], but the premise is that supervision institutions have a full understanding of power enterprises. In literature [15], an identification method using the fuzzy estimator to evaluate market power is introduced.

However, the following problems still exist when applying these methods to China's power market. First of all, the foreign electricity market has been open for a long time, so that their day-ahead and spot transactions have been relatively mature and occupy the majority of the market trading volume. But the reform of the domestic electricity market has just begun, it is still dominated by the medium-term and long-term transactions. So it is necessary to design China’s market power violation identification method for China’s trading rules. Secondly, due to the high marketization level of foreign power markets, most of the supervision methods are macro supervision for the whole market. But in China, due to the lack of competition in the early stage, it is impossible to rely on the market itself to adjust, so it is more urgent to identify whether a market subject has abused its market power. But there are few ways to identify it, so new methods of identification need to be studied.

The structure of the paper is as follows. In the first section, to form an identification system, this paper analyzes market power indexes of the whole market and individual market subject, summarizes the advantages and disadvantages of them. Although the index identification system covers the principle, and manual judgment is accurate, they are less efficient. Therefore in the second section, this paper introduces an intelligent recognition algorithm – support vector machine(SVM), and improved based on the data imbalance problem. To take into account the economic principles of market power as well as the need for rapid identification in the meantime, this paper combines manual judgment with intelligent recognition, use index identification system to assist the artificial judgment and construct the training set, which is then trained and tested by intelligent recognition algorithm. Then the third section takes the simulation data as an example. The experimental results show that the violation recognition rate reached 100%, which reflects the algorithm is feasible and effective. Finally, the fourth part carries on the summary and the forecast to the full text.

2. Identification index system

2.1 Indexes to measure market power of the whole market

Indexes of the whole market can be divided into two types: pre-bidding indexes and post-bidding indexes. The pre-bidding index measures the degree of market concentration, which to some extent reflects the level of market power, mainly including Top-m Share Index and Herfindahl-Hirschman Index (HHI). The post-bidding index is used to measure the degree the transaction price deviating from the marginal production cost of the market, mainly including the Lerner index (LI).

2.1.1 Pre-bidding index (Top-m Share Index and HHI) Top-m Share Index refers to the quotient of electric bidding capacity of the largest m power enterprises and that of the whole market. The higher the value is, the higher the market concentration degree is. The Top-4 index is usually adopted. When it is higher than 65%, it indicates the market is oligopolistic and the market concentration is high.

Top-m Share Index’s principle is easy to calculate, which to a certain extent reflects the level of the existence of market power, but it ignores information of most enterprises, so its information utilization degree is poor, the HHI index is an improvement of it.

HHI is the sum of squares of the market shares of each generation enterprise. The formula is:

$$\text{HHI} = \sum_{i=1}^{n} \left( 100 \times \frac{q_i}{q_{\text{sum}}} \right)^2,$$

where $n$ is the number of power enterprises and $q_i$ is the bidding capacity of the $i$th power enterprise. It is considered that the market with the HHI value less than 1800 has sufficient competition.
2.1.2 Post-bidding index (LI) The Lerner Index is the degree of the transaction price deviates from the marginal production cost of the market, and the calculation formula is as follows:

\[ LI = \frac{p - Mc}{p}. \]  

(2)

where \( p \) is the transaction price and \( Mc \) is the marginal production cost. It is easy to see that the higher the LI value is, the more the transaction price deviates from the marginal production cost.

But such price anomaly may not only be caused by market power. Therefore, the Lerner Index can only reflect the effect of market power to some extent. The more difficult question is how to calculate the marginal production cost. The market bidding rules make sure the uniform clearing pricing is equal to the marginal production cost of the whole market. However, the domestic power market is still dominated by the medium-term and long-term transactions. It is only divided into three times every day, each paragraph only can be quoted one or three different prices. In this case, the marginal price can not reflect the marginal production cost, resulting in the LI index unusable.

2.2 Indexes to measure market power of an individual market subject

Indexes of an individual market subject can be divided into three types: pre-bidding index, in-bidding index, and post-bidding index. Pre-bidding indexes are used to measure whether a power enterprise has market power, which mainly include the Market Share (MS), the Must-run Ratio (MRR) and the MRR considering transmission restriction (i.e. improved IMRR). The in-bidding index is used to measure whether an enterprise has the intention to use market power. The main indexes are the High Price Bidding Ratio (HPBR), and Retention Rate (RR). The post-bidding index measures the affect of the market power, which can be calculated by the Lerner Index of an individual enterprise (LII) and Price Difference (PD). These indexes are introduced as follows:

2.2.1 Pre-bidding index (MS, MRR, and IMMR) Market Share is a direct reflection of an enterprise’s scale, and the calculation formula is as follows:

\[ s_i = \frac{q_i}{\sum_{i=1}^{n} q_i}, \]

(3)

where \( n \) is the number of power enterprises, \( q_i \) is the bidding capacity of enterprise \( i \). However, as the electricity market is basically in a state of oversupply, even enterprises with large MS value do not necessarily have great market power, which is also related to the supply and demand relationship of the market. The following index MRR can take it into consideration.

MRR refers to the ratio of the market power that a plant must generate to meet the system load demand to the power that a plant can generate in a certain period. The calculation formula is as follows:

\[ MRR_j = \max \left\{ \frac{D - \sum_{l=1}^{n} q_j}{C_i}, 0 \right\}, \]

(4)

where \( D \) represents the total market demand, \( q_j \) is the bidding capacity of the power enterprise \( j \), and \( C_i \) represents the power that the power enterprise \( i \) can generate.

When MMR is not equal to zero, the power enterprise has absolute market power. Although MMR considers the supply and demand relationship, it can only represent the absolute market power, which is only owned with few power enterprises. Besides, the relationship between enterprise size and market power is not absolute, and market power is also related to transmission restrictions.

To calculate IMMR, the way to calculate the must-run power capacity considering transmission restriction is necessary. The calculation formula is as follows:
\[
\begin{align*}
\min P_{R_k}. \\
\text{Such that} \\
e^T = (P_g - P_d) = 0, \\
0 \leq P_g \leq P_{g_{\text{max}}}, \\
-PL_{\text{max}} \leq F(P_g - P_d) \leq PL_{\text{max}},
\end{align*}
\]

where \( e \) is a vector with all ones, \( P_g \) is the power dispatch vector, \( P_d \) is the demand vector, \( PL_{\text{max}} \) is the line limit vector and \( F \) is the distribution factors matrix.\(^{[5, 7]}\) Set the calculation result of the above equation as \( P_{g_{\text{must}}} \), the IMMR can be calculated as:

\[
MRS = \frac{P_{g_{\text{must}}}}{P_d},
\]

This index takes into account the influence of transmission congestion. Although it can truly reflect the market power of the enterprise, it is complicated to calculate. Moreover, it needs to consider the load and scheduling of each line of the entire network, which is difficult for the independent trading center.

### 2.2.2 In-bidding index (HPBR and RR)

HPBR can be calculated as follows:

\[
\text{HPBR}_i = 100\% \times \frac{q_i^{\text{high}}}{q_i},
\]

where \( q_i \) represents total bidding capacity of power enterprise \( i \), and \( q_i^{\text{high}} \) is the bidding capacity belongs to high price.

But the definition of high price is controversial. In Beijing trading center, it is the price that is higher than the average of all power enterprise’s bidding prices plus \( \max \{0.03, 1.6 \times \sigma \} \) yuan, where \( \sigma \) is the standard deviation of bidding price. Whereas, Guangdong defines it as the last 5% of the quote curve. However, both definitions hasn’t consider the economic principles. Actually, the definition should be based on the marginal cost of the enterprise itself, which is difficult to know in China’s market, as we discussed before. But as the index is essential to reflect the intentions of market power using, it can be adopted pool purchase price as the marginal cost. Thus, this paper defines it as a price higher than the marginal cost of power generation \( \times (1 + \times) \), where \( x \) should be set according to the market situation.

Power enterprises may use their market power to raise prices by maliciously reducing power generation, which can be reflected by RR. The calculation formula is as follows:

\[
\text{RR}_i = 100\% \times \frac{q_i^{\text{high}}}{C_i},
\]

where \( q_i \) represents the total bidding capacity of power enterprise \( i \), \( C_i \) represents the power that the power enterprise \( i \) can generate. The biggest problem with this index is that it’s difficult to obtain \( C_i \).

### 2.2.3 Post-bidding index (LII and PD)

The definition and calculation formula of LII is similar to L1. However, market power is not the only reason for its anomaly. The following index PD can avoid this problem. Whereas, PD refers to the difference between the actual transaction price and the transaction price when power enterprise i don't use its market power (i.e., it bids as its marginal production cost). It can well reflect the market price deviation caused by the use of market power.

### 2.3 Index system

To construct an identification index system of market power abuse, combining indexes of the whole market and individual market subject is needed. Firstly, using the pre-bidding index of the whole market to judge the degree of market concentration and determine the intensity of supervision. Then, comprehensively consider the pre-bidding, in-bidding, post-bidding indexes of each power enterprise to make the judgment that whether they have violated the rules. Considering the difficulty of information acquisition and calculation, the index system is constructed and shown in the figure 1.
3. Intelligent identification method

Violation identification is a classification problem, and the data of power enterprises can be divided into two types: violations and non-violations. Support Vector Machine (SVM) is a classification method with solid mathematical theory. The SVM has a good generalization ability. Therefore, this paper chooses SVM for violation identification.

3.1 Principle of SVM

For linearly separable data, the principle of the support vector machine is as follows. Set a training set as \( \{ (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \} \), \( y_i \in \{1, -1\}, \quad i = 1, 2, \ldots, n \), where \( n \) is the number of power enterprises; \( x_i \) is bidding data. For example, in a three-part bidding rule, \( x_i = (x_{i1}^p, x_{i2}^p, x_{i3}^p) \), where \( x_{ij}^p \) respectively indicate the bidding price and bidding capacity of the \( j \)th stage of the enterprise \( i \), \( j = 1, 2, 3 \); \( y_i \) are the results of identification, of which 1 represents non-violation data, and -1 is violation data. Set the optimal classification interface as:

\[
 w^T x + b = 0.
\]  

Then the mathematical model of the nearly linear SVM is:

\[
 \min_{w} \quad f(w) = \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i,
\]

subject to:

\[
 y_i (w^T x + b) \geq 1 - \xi_i,
\]

\[
 \xi_i \geq 0.
\]

where \( \xi_i \) is the slack variable, \( C \) is the penalty parameter, the larger the \( C \) is, the greater the penalty for misclassification. According to the Lagrangian dual principle, (10) can be converted into:

\[
 \max_{\lambda} \quad \left( \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_i \lambda_j y_i y_j x_i^T x_j \right),
\]

subject to:

\[
 \sum_{i=1}^{n} \lambda_i y_i = 0,
\]

\[
 C \geq \lambda_i \geq 0.
\]

(11) is a solvable quadratic programming problem. Set the solution is \( \lambda^* \), then:

\[
 w^* = \sum_{i=1}^{n} \lambda_i^* y_i x_i, \quad b^* = y_i - w^T x_i.
\]

The decision function is:

\[
 f(x) = \text{sgn} \left[ w^T x + b \right] = \text{sgn} \left[ \sum \lambda_i^* y_i x_i^T x + b^* \right].
\]

When the train set is linearly inseparable, the kernel function needs to be introduced to map the data from the low-dimensional space to the high-dimensional space, converting it into linearly separable data.
It is defined as:

\[
K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle = \phi^T(x_i) \cdot \phi(x_j),
\]

where \( \langle \cdot, \cdot \rangle \) is the inner product.

3.2 Improved SVM

In practice, the proportion of data identified as violations is very small, so the training set is faced with the problem of unbalanced data, making the bad classification result. Therefore, the Synthetic Minority Oversampling Technique (SMOTE) is used to increase the number of violation data. The specific algorithm is as follows:

- **Step 1**: Setting an oversampling ratio \( N \) according to the sample imbalance ratio;
- **Step 2**: Randomly sample a violation data \( x \) from the training set, calculate its Euclidean distance to all samples in the violation data set \( V \), and obtain its \( k \)-nearest neighbor;
- **Step 3**: Randomly select \( N \) samples from its \( k \)-nearest neighbors, then construct a new sample for each neighbor \( x \) according to the following formula:

\[
x_{\text{new}} = x + \text{rand}(0,1) \times |x - x_{i}|,
\]

where \( \text{rand}(0,1) \) is a random number between 0 and 1;
- **Step 4**: If the two types of samples are equal in number, the algorithm ends, otherwise return to Step 2.

3.3 Violation identification algorithm flow

Through the above discussion, the algorithm flow of this paper is as follows:

- **Step 1**: Calculate the indexes of index system using the original transaction data, and make a manual judgment to construct a training set;
- **Step 2**: Use the SMOTE algorithm to extend the data set so that the number of violating samples is equal to the number of non-violated samples;
- **Step 3**: Standardize the expanded training data set;
- **Step 4**: Using SVM for training;
- **Step 5**: Data testing and results analysis.

3.4 Algorithm evaluation criteria

The training set can be divided into two types: violation data \( n_v \) and non-violation data \( n_n \). The results of the classification algorithm can be divided into four categories: the violation data which is identified as the violation data \( n_{vv} \), the violation data which is identified as the non-violation data \( n_{vn} \), the non-violation data which is identified as the non-violation data \( n_{nn} \), and the non-violation data which is identified as the violation data \( n_{nv} \). The algorithm can be evaluated with the correct rate and recall rate.

The formula is as follows:

\[
\text{Correct rate} = \frac{n_{vv} + n_{nn}}{n_v + n_n}, \quad \text{Recall rate} = \frac{n_{vv}}{n_v}.
\]

The correct rate can reflect the overall classification accuracy level. The recall rate can reflect whether the violation data is fully identified, and ignore the correct rate of non-violation data. An algorithm with a high correct rate on the premise that the violation data is fully recognized is what we need.

4. Example

In this paper, the data of 42 power enterprises in a simulated transaction are used for example analysis, and part of the data are shown in Table 1.

The clearing price of this transaction is 304.32 yuan/MWh, and the HHI value calculated by the formula (1) is 1328.12. Although it is less than 1800, many power enterprises are controlled by state-owned enterprises, so the supervision cannot be eased. Therefore, the high price is defined as the
price that is higher than the enterprise's marginal production cost \times (1 + 30\%). In this paper, the marginal production cost is calculated by the Monte Carlo method, and the specific method is shown in the literature [16].

Table 1. Original transaction data of some power enterprises.

| Generate-on enterprise | First capacity | First price | First capacity | Second price | Second capacity | Third price | Third capacity | Marginal generate-on cost |
|------------------------|---------------|------------|---------------|--------------|----------------|------------|----------------|-------------------------|
| 3                      | 47361         | 298.01     | 47361         | 16412        | 301.67         | 1227       | 303.63         | 1227                    |
| 9                      | 51023         | 298.20     | 51023         | 25511        | 301.62         | 15959      | 304.82         | 232.24                  |
| 31                     | 486361        | 298.00     | 486361        | 7975         | 299.01         | 7975       | 5662           | 8465                    |
| 33                     | 553779        | 298.02     | 553779        | 15959        | 304.28         | 15959      | 306.42         | 231.25                  |
| 34                     | 453270        | 299.00     | 453270        | 97594        | 303.64         | 97594      | 99135          | 231.30                  |

4.1 Using the identification index system to construct the training set
Firstly, this paper use the original data to calculate the indicators in the index system and make manual judgments. Partial results are shown in Table 2.

4.2 Oversampling process and data standardization
Part of the results after the sampling process and data standardization are shown in Table 3.

Table 2. Index data of some power enterprises.

| Power enterprises | MS         | HPBR       | PD /yuan | Judgement |
|-------------------|------------|------------|----------|-----------|
| 3                 | 1.20%      | 0          | 0        | 1         |
| 9                 | 1.57%      | 40.00%     | 0        | -1        |
| 31                | 9.23%      | 1.13%      | 0.17     | 1         |
| 33                | 11.08%     | 7.70%      | 2.18     | -1        |
| 34                | 12.00%     | 30.27%     | 3.12     | -1        |

Table 3. Data after standardization.

| Power enterprises | MS         | HPBR       | PD /yuan | Judgement |
|-------------------|------------|------------|----------|-----------|
| 3                 | 0.1        | 0          | 0        | 1         |
| 9                 | 0.13       | 1          | 0        | -1        |
| 31                | 0.77       | 0.03       | 0.05     | 1         |
| 33                | 0.92       | 0.19       | 0.70     | -1        |
| 34                | 1          | 0.76       | 1        | -1        |

4.3 Support vector machine training and result analysis
Using SVM for training by MATLAB, the training time is 0.43 seconds. Still using this 42 data for testing, the test time is 0.04 seconds. And the results are shown in Table 4.

Table 4. Test correct rate and recall rate of SVM.

| Index            | Linear SVM | Gaussian kernel function | Polynomial kernel function | Sigmoid kernel function |
|------------------|------------|--------------------------|----------------------------|------------------------|
| Correct rate     | 89.74%     | 92.31%                   | 100%                       | 61.54%                 |
| Recall rate      | 100%       | 100%                     | 100%                       | 100%                   |

It can be seen from Table 4 that the linear inseparable support vector machine with the polynomial function as the kernel function has the highest correct rate under the premise of ensuring that all the violation data is recognized. When the polynomial with exponent of 3 is chosen as the kernel function, the correct rate and recall rate are both 100%, so the algorithm is feasible and effective.

5. Conclusion
This paper focuses on the application of the identification of market power abuse in China’s power market. Two problems were found and analyzed in the application process, that is, the trading rules and marketization degree which are different from the foreign market. And an index system as well as an intelligent algorithm are proposed. The main conclusions of this paper are:

(1) The identification of market power abuse in China’s power market must conform to China’s actual situation and be adjusted according to China’s trading rules and marketization degree;

(2) The data set of the SVM algorithm is the index data constructed by the index system, which realizes the combination of the traditional identification index and the intelligent algorithm, which can
well explore the potential relationship of index data.

(3) The occurrence of violations reflects the deficiency of trading rules to some extent. In the following work, it is necessary to modify trading rules from the analysis of violations so as to avoid the occurrence of violations to the greatest extent.

Acknowledgment
This paper was supported by the Science and Technology Projects of State Grid Co., Ltd. (Grant No. SGGSJY00PSJS1900060).

References
[1] Shi, S.H., Liu, D.N., Hu, H.W., et al. (2018) Research on the identification of irregularities in the whole process of power trading under the open market. Price: theory&practice, 08: 51-54. (in Chinese)

[2] Xia, Q., Li, C.B., Jiang, J.J., et al. (2003) Methods, indicators and means of foreign power market supervision[J]. Power System Technology, 27(3): 1-4. (in Chinese)

[3] Karthikeyan, S.P., Raglend, I.J., Kothari, D.P. (2013) A review on market power in deregulated electricity market. JEPE, 48(1):139–147.

[4] Lakic, E., Medved, T., Zupancic, J., et al. (2017) The review of market power detection tools in organised electricity markets. In: EEM. Dresden, Germany.

[5] Yu, Z., Nderitu, D.G., Sparrow, F.T. (2000) A proposed market power monitoring model for restructuring electricity markets. IEEE Power Engineering Review, 20(7):41-42.

[6] Bompad, E., Ma, Y.C., Napoli, R., et al. (2006) Assessing the market power due to the network constraints in competitive electricity markets. Electric Power Systems Research, 76(11):953-961.

[7] Rahimi, F., Zhang, M., Hobbs, B. (2007) Competitive Transmission Path Assessment for Local Market Power Mitigation. In: Power Engineering Society General Meeting. Tampa, Florida.

[8] Talebazloo, B.F., Latify, M.A., Rahimiyan, M. (2015) Ex-ante analysis of potential impact of transmission switching on market power. In: Electrical Engineering. Tehran, Iran. pp:1580-1584.

[9] Shukla, U.K., Thampy, A. (2011) Analysis of competition and market power in the wholesale electricity market in India. Energy Policy, 39(5):2699-2710.

[10] Ruksans, O., Oleinikova, I., Junghans, G. (2014) Electricity market integration and monitoring in Baltics. In: EEM14. Krakow, Poland.

[11] Chakrabarti, B.B., Goodwin, D.G. (2008) Monitoring and Measuring Market Power in the New Zealand Electricity Market. In: 2008 Joint International Conference on Power System Technology and IEEE Power India Conference. New Delhi, India.

[12] Zhang, L.Z. (2017) How does UK electricity reform regulate market power. China Energy News,2017-11-06(004). (in Chinese)

[13] Khajeh, H., Foroud, A.A., Firoozi, H. (2016) New ex-ante indices of market power. In: ICEE. Shiraz, Iran. pp: 67-71

[14] Soleymani, S., Ranjbar, A.M., Jafari, A., et al. (2006) Market power monitoring in electricity market by using market simulation. In: IEEE Power India Conference. New Delhi, India.

[15] Mashhadi, R., Habib, Rahimiyan, et al. (2011) Measurement of Power Supplier’s Market Power Using a Proposed Fuzzy Estimator. IEEE Transactions on Power Systems, 26(4):1836-1844.

[16] Zhang, X.S., Guo, X.L., Yi, L.F. (2011) Marginal cost analysis of electricity generation based on monte carlo simulation. Power System Protection and Control, 39(14): 21-25. (in Chinese)